MATHEMATISCHE ANNALEN

BEGRÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN

FORTGEFÜHRT DURCH

FELIX KLEIN

GEGENWÄRTIG HERAUSGEGEBEN
VON

DAVID HILBERT
IN GÖTTINGEN

UNTER MITWIRKUNG VON

OTTO BLUMENTHAL
IN AACHEN

ERICH HECKE
IN HAMBURG

102. BAND

BERLIN
VERLAG VON JULIUS SPRINGER
1930

MATHEMATISCHE ANNALEN

BEGRÜNDET 1868 DURCH

ALFRED CLEBSCH UND CARL NEUMANN

FORTGEFÜHRT DURCH

FELIX KLEIN

GEGENWÄRTIG HERAUSGEGEBEN

VON

DAVID HILBERT

IN GÖTTINGEN

UNTER MITWIRKUNG VON

OTTO BLUMENTHAL
IN AACHEN

ERICH HECKE
IN HAMBURG

102. BAND

BERLIN
VERLAG VON JULIUS SPRINGER
1930

Inhalt des einhundertundzweiten Bandes.

(In alphabetischer Ordnung.)

Probleme der Grundlegung der Mathematik[1]).

Von

David Hilbert in Göttingen.

Für die mathematische Wissenschaft waren die letzten Jahrzehnte eine Periode höchster Blüte.

Ich erinnere daran, wie in der Arithmetik, insbesondere in der Theorie der algebraischen Zahlkörper, die schwierigsten Probleme gelöst wurden und die Vollendung dieses herrlichen Gedankengebäudes erreicht worden ist. Zugleich gelang die Entdeckung der lange gesuchten transzendenten Funktionen, die dem Zahlkörper zugehören, und durch die mannigfache bisher verborgene zahlentheoretische Wahrheiten ans Licht gelangten. Andererseits wurden die Begriffsbildungen der Idealtheorie weit über die Grenzen der Zahlentheorie hinaus auf Algebra und Funktionentheorie mit bestem Erfolge übertragen und dadurch ein großer Komplex mathematischen Wissens zu einem einheitlichen Gefüge gemacht.

Auch in der Theorie der Funktionen einer komplexen Veränderlichen sind in dem verflossenen Zeitraum nicht geringe Fortschritte gemacht worden. Der Ausbildung des Prinzips der konformen Abbildung insbesondere verdanken wir die herrlichen Methoden zur Gewinnung der automorphen Funktionen und die Lösung des Problems der Uniformisierung. Die so schwierigen Beweise der Existenzsätze haben durch Anwendung der Methoden der Variationsrechnung den höchsten Grad von Einfachheit und Durchsichtigkeit erreicht.

Und welche Fülle der Gesichte liefert die Geometrie: Allein die Topologie ist so sehr durch neue Fragestellungen und Behandlungsmethoden bereichert worden, daß man darin die Entstehung eines neuen selbständigen Wissenszweiges sehen muß. Und die alten der Geometrie nahestehenden Disziplinen, die Gruppen- und Invariantentheorie, sind zu einer ungeahnten Erweiterung und Vertiefung gelangt.

---

[1]) Vortrag, gehalten auf dem Internationalen Mathematiker-Kongreß in Bologna am 3. Sept. 1928.

Die Physik endlich ließ vor unseren Augen mathematische Gedankengebäude erstehen, deren Hallen von imponierender Großartigkeit sind. Überhaupt gedenken wir auch der Anwendungen: es sind nicht die schlechtesten Früchte, die die mathematische Forschung auf dem Felde der Anwendungen erntet, sei es, daß die Anwendungen benachbarten Wissensgebieten oder praktischen Bedürfnissen entsprungen sind. Und der Bezirk, in den die Mathematik eindringt, dehnt sich beständig.

Wegen dieser so erfreulichen Sachlage erwächst dem Mathematiker die besondere Pflicht, die Mathematik in ihren Fundamenten zu sichern.

Wie ist doch die allgemein verbreitete Meinung über Mathematik und mathematisches Denken? Sie lautet: Die mathematischen Wahrheiten sind absolut sicher, denn sie werden auf Grund von Definitionen durch untrügliche Schlüsse bewiesen. Sie müssen daher auch überall in der Wirklichkeit stimmen. Dieser populären Meinung zufolge soll die Mathematik zum Vorbild für alle Wissenschaft überhaupt dienen. Wir wollen nun sehen, ob es so mit der Mathematik bestellt ist. Wie war der Stand der Grundlagenfrage im drittletzten Jahrzehnt? Die großen Klassiker und Schöpfer der Grundlagenforschung waren Cantor, Frege und Dedekind; sie hatten in Zermelo einen kongenialen Interpreten gefunden. Zermelo hatte die Annahmen, die zum axiomatischen Aufbau der Mengenlehre nötig sind, aufgestellt und damit die von Cantor und Dedekind nur unbestimmt und teilweise unbewußt angewandten Mittel präzisiert. Diese Zermeloschen Axiome sind zudem alle von der Art, daß ein ernster Zweifel an ihrer Richtigkeit nicht aufkommen konnte. Das Vorgehen Zermelos war durchaus berechtigt und entspricht der axiomatischen Methode. Doch die Zermeloschen Bahnen wurden unter dem Druck maßgebender Kreise verlassen. Alte Einwürfe Kroneckers gegen Cantor und Dedekind, die wir längst überwunden glaubten und die Kronecker selbst in seinen Arbeiten nicht befolgt hatte, wurden vorgesucht. Eine unglückliche Auffassung Poincarés, betreffend den Schluß von n auf $n+1$, die bereits zwei Jahrzehnte früher von Dedekind durch einen präzisen Beweis widerlegt worden war, verrammelte den Weg zum Vorwärtsschreiten. Ein neues Verbot, das Verbot der imprädikativen Aussagen, wurde von Poincaré erlassen und aufrechterhalten, obwohl Zermelo sofort ein schlagendes Beispiel gegen dieses Verbot angab und dieses Verbot außerdem gegen die Resultate Dedekinds verstieß. Leider auch die sonst so vortreffliche Logik Russells leistete in ihrer Anwendung auf Mathematik der Irrlehre Vorschub. So kam es, daß unsere geliebte Wissenschaft — was die Frage nach ihrem innersten arithmetischen Wesen und Grunde betrifft — zwei Jahrzehnte hindurch wie von einem bösen Traum heimgesucht wurde.

Ich begrüße es als ein Erwachen, als ein leuchtendes Morgenrot, wenn in der letzten Zeit eine Reihe jüngerer Mathematiker wieder auf Zermelos Gedanken zurückgehen; diese Mathematiker haben die Zermeloschen Axiome vervollständigt und dazu eine Reihe sehr wichtiger tiefliegender Fragen erfolgreich behandelt.

Freilich eine endgültige Lösung der Grundlagenprobleme ist durch dieses axiomatische Verfahren niemals möglich. Denn die von Zermelo zugrunde gelegten Axiome enthalten echte inhaltliche Annahmen, und diese zu beweisen ist, wie ich glaube, gerade die Hauptaufgabe der Grundlagenforschung, — war doch schon damals die Widerspruchsfreiheit der arithmetischen Axiome zu beweisen eine brennende Frage geworden. Wenn wir aber inhaltliche Axiome als Ausgangspunkte und Grundlagen für die Beweise benutzen, so verliert die Mathematik damit den Charakter des absolut Sicheren. Mit der Annahme von Voraussetzungen kommen wir auf das Gebiet des Problematischen, — beruhen doch die Meinungsverschiedenheiten der Menschen meist darauf, daß von verschiedenen Voraussetzungen ausgegangen wird. In einer Reihe von Vorträgen im Laufe der letzten Jahre habe ich daher einen neuen Weg zur Behandlung des Grundlagenproblems betreten. Mit dieser Neubegründung der Mathematik, die man füglich als eine Beweistheorie bezeichnen kann, glaube ich die Grundlagenfragen in der Mathematik als solche endgültig aus der Welt zu schaffen, indem ich jede mathematische Aussage zu einer konkret aufweisbaren und streng ableitbaren Formel mache und dadurch den ganzen Fragenkomplex in die Domäne der reinen Mathematik versetze.

Freilich bedarf es zur völligen Durchführung dieser Aufgabe der hingebenden Mitarbeit der jüngeren Mathematikergeneration.

In diesem Sinne möchte ich heute einige nähere Ausführungen machen. Die wichtigsten Probleme konzentrieren sich sämtlich auf das von mir aufgestellte sogenannte ε-Axiom; dasselbe lautet

$$A(a) \rightarrow A(\varepsilon A).$$

Bei der Verwendung des Axioms hat man vor allem die Gattung von Variablen zu beachten, auf die ε zu beziehen ist. Bei den Zahlen dient dasselbe zur Formulierung der üblichen Schlüsse mit „irgendein“: Man versteht unter $\varepsilon\mathfrak{A}$ irgendeine Zahl, für die die Aussage $\mathfrak{A}$ sicher zutrifft, wenn es überhaupt eine Zahl gibt, für die $\mathfrak{A}$ zutrifft.

Ich möchte nun einige Probleme nennen.

Durch die Arbeiten von Ackermann und v. Neumann ist der Beweis für die Widerspruchsfreiheit des ε-Axioms für die Zahlen geführt; damit sind folgende drei Probleme erledigt:

1. Das Tertium non datur für die Zahlen, d. h. wenn eine Aussage nicht für alle ganzen Zahlen gilt, so gibt es eine Zahl, für die sie nicht gilt.

Beispielsweise: nach Kronecker war es unzulässig, eine ganze rationale Funktion einer Variablen x mit ganzen rationalen Koeffizienten als irreduzibel zu definieren, wenn es keine Zerlegung derselben als Produkt von zwei ebensolchen Funktionen gibt. Ich beweise durch die Beweistheorie, daß im Gegenteil diese Definition eine völlig strenge Definition in rein mathematischem Sinne ist, und daher war die Behauptung Kroneckers nicht bloß logisch unzutreffend, sondern in rein arithmetischem Sinne unrichtig — in gleichem Sinne unrichtig wie irgendein falscher arithmetischer Satz oder eine falsche zahlentheoretische Formel.

2. Die Auflösung einer Behauptung über Existenz einer Zahl nach dieser Zahl: „die kleinste Zahl, welche“.

3. Der Schluß von n auf $n+1$, wenn man noch die Formel

$$(\varepsilon A = b') \rightarrow \bar{A} b$$

als Axiom hinzunimmt.

Wie Sie bemerken, ist ein wesentliches Hilfsmittel für meine Beweistheorie die Begriffsschrift; wir verdanken dem Klassiker dieser Begriffsschrift, Peano, die sorgfältigste Pflege und weitgehendste Ausbildung derselben. Die Form, in der ich die Begriffsschrift brauche, ist wesentlich diejenige, die Russell zuerst eingeführt hat.

Bisher noch nicht erledigt sind die folgenden Probleme:

Problem I.

Der Nachweis der Widerspruchsfreiheit des ε-Axioms für die Funktionsvariable f. Der Ansatz eines Beweises liegt vor. Diesen hat Ackermann schon so weit durchgeführt, daß die verbleibende Aufgabe nur noch in dem Beweise eines rein arithmetischen elementaren Endlichkeitssatzes besteht.

Mit der Lösung dieses Problems I ist bereits ein sehr großer Komplex von grundsätzlichen Fragen behoben; nämlich dieser Nachweis der Widerspruchsfreiheit liefert:

1. das Tertium non datur für Funktionen von ganzen Zahlen und damit auch für die reellen Zahlen;

2. die Definitionsprozesse, die von Poincaré als imprädikativ angefochten worden sind, die Russell und Whitehead nur mit Hilfe des sehr problematischen Reduzierbarkeitsaxioms begründen konnten, und mit Bezug auf die Weyl einmal von einem Circulus vitiosus in der Analysis gesprochen hat;

3. das Auswahlaxiom in der schwächeren Form.

Zu 3. sei folgendes bemerkt. Es werden in den neueren Betrachtungen über das Auswahlaxiom vielfach Unterscheidungen gemacht zwischen schwächeren und schärferen Formen des Auswahlprinzips. Diese betreffen zumeist die Mächtigkeit der Mengen von Mengen und deren Repräsentantenmenge, insbesondere den Unterschied des Abzählbaren und Überzählbaren.

Durch die Beweistheorie werden wir veranlaßt, vor allem eine anderweitige Unterscheidung als wesentlich anzusehen, nämlich, ob verlangt wird, daß die Auswahl des Repräsentanten für eine Menge, unabhängig von der Art, wie die Menge definiert ist, eindeutig durch den Elementenbestand der Menge bestimmt sein soll, oder ob dies nicht verlangt wird.

Beispielsweise sei eine einparametrige Mengenschar $M(t)$ vorgelegt. Für jeden Wert des reellen Parameters t bezeichne $M(t)$ eine bestimmte Menge von reellen Zahlen, die mindestens ein Element enthält.

Das Auswahlprinzip in der schwächeren Form verlangt dann, daß es eine eindeutige Funktion $f(t)$ gibt von der Art, daß für jeden Wert von t der Wert $f(t)$ zu $M(t)$ gehört. In der stärkeren Form verlangt das Auswahlprinzip überdies die Existenz einer solchen Funktion $f(t)$, daß

$$f(t_1) = f(t_2)$$

ist, falls die Mengen $M(t_1)$ und $M(t_2)$ ihrem Elementenbestande nach übereinstimmen.

Mit Hilfe des ε-Axioms für die Funktionenvariable f erhalten wir das Auswahlprinzip für Mengen von reellen Zahlen in der schwächeren Form.

Durch die Lösung unseres Problems I werden vor allem folgende Theorien beherrscht:

1. Die Theorie der reellen Zahlen. (Dedekindscher Schnitt, obere Grenze einer beschränkten Menge reeller Zahlen.)

2. Die Peanosche Begründung der Zahlenlehre. — In dieser Theorie braucht man kein Auswahlprinzip, wohl aber die imprädikative Definition, z. B. die Definition für $a \leq b$: nämlich, jede Menge, die a enthält und mit n zugleich $n+1$, enthält auch b. Man hat früher das Problematische bei den mengentheoretischen Begründungen der Zahlentheorie immer nur in der Voraussetzung eines unendlichen Individuenbereichs gesehen. Diese Voraussetzung kann bereits auf Grund des Bisherigen als widerspruchsfrei erkannt werden. Die größere Schwierigkeit liegt in dem Nachweis der Widerspruchsfreiheit der imprädikativen Definition.

Mit der Lösung des Problems I erhält auch das geniale Verfahren Dedekinds in dessen Abhandlung „Was sind, was sollen die Zahlen?“ die Rechtfertigung.

3. Die Cantorsche Theorie der Wohlordnungen der Zahlenreihe. Durch diese wird die Lehre von den Zahlen der zweiten Zahlenklasse, welche sich axiomatisch ganz analog zu der Zahlentheorie aufbauen läßt, auf die Theorie der Funktionen von Zahlenvariablen zurückgeführt.

Problem II.

Für die weitere Ausgestaltung der Analysis, insbesondere für die Punktmengen-Theorie (mengentheoretische Topologie) sowie auch für die Theorie der zweiten und der höheren Zahlklassen, hat man die Widerspruchsfreiheit für das ε-Axiom, bezogen auf höhere Variablengattungen, zunächst einmal auf die der Funktionen reeller Variablen nötig.

Ferner gebraucht man zum Beweis des Wohlordnungssatzes und auch für manche Beweise in der Theorie der Meßbarkeit (Beweise für Nichtmeßbarkeit) das Auswahlaxiom in der stärkeren Form, welches in der Beweistheorie durch die Formel

$$((f)(A(f) \rightleftarrows B(f))) \rightarrow \varepsilon A = \varepsilon B$$

ausgedrückt wird (Axiom der Auswahlgleichheit); εA und εB sind hier Funktionen der Zahlenvariablen und die Gleichheit bedeutet Übereinstimmung für alle Argumentwerte. Für die Hinzunahme dieser Formel ist wiederum der Beweis der Widerspruchsfreiheit nötig.

Problem III.

Die Vollständigkeit des Axiomensystems für die Zahlentheorie sowie für die Analysis wird zwar allgemein behauptet; aber die übliche Überlegung, mit der man zeigt, daß je zwei Realisierungen des Axiomensystems der Zahlentheorie bzw. der Analysis isomorph sein müssen, genügt nicht den Anforderungen finiter Strenge.

Es kommt darauf an, zunächst für die Zahlentheorie, deren Bereich sich präzise abgrenzen läßt, den üblichen Beweis der Isomorphie finit umzugestalten, so daß dadurch folgendes gezeigt wird:

Wenn für einen Satz $\mathfrak{S}$ die Widerspruchsfreiheit mit den Axiomen der Zahlentheorie nachgewiesen werden kann, so kann nicht auch für $\overline{\mathfrak{S}}$ (das Gegenteil von $\mathfrak{S}$) die Widerspruchsfreiheit mit jenen Axiomen nachgewiesen werden.

Und damit in engstem Zusammenhang auch: wenn eine Aussage widerspruchsfrei ist, so ist sie auch beweisbar.

In höheren Gebieten wäre der Fall der Widerspruchsfreiheit von $\mathfrak{S}$ sowie der von $\overline{\mathfrak{S}}$ denkbar: alsdann ist die Annahme einer der beiden Aussagen $\mathfrak{S}$, $\overline{\mathfrak{S}}$ als Axiom durch systematische Vorzüge (Prinzip der Permanenz der Gesetze, weitere Aufbaumöglichkeiten usw.) zu rechtfertigen.

Gegen meine Beweistheorie sind Einwände erhoben worden; diese beruhen wesentlich auf einer Verkennung meiner Beweistheorie. Ich gestatte mir daher, hier noch einige Erläuterungen zu machen.

Wenn eine Aussage oder ein Beweis vorliegt, so muß er sich in allen Teilen überblicken lassen. Die Aufweisung, das Wiedererkennen, die Unterscheidung und Aufeinanderfolge seiner einzelnen Teile ist unmittelbar anschaulich für uns da. Ohne diese Einstellung ist überhaupt ein Denken oder gar eine wissenschaftliche Tätigkeit unmöglich. Bei der Untersuchung der Widerspruchsfreiheit handelt es sich darum, ob ein Beweis vorgelegt werden kann, der zu einem Widerspruch führt. Wenn mir ein solcher Beweis nicht vorgelegt werden kann, um so besser, — da mir dann ein Eingehen erspart bleibt. Wenn mir aber der Beweis vorliegt, so darf ich gewisse einzelne Teile herausgreifen und für sich ins Auge fassen, also auch die besonderen Zahlzeichen, welche in ihnen vorkommen und also hergestellt und aufgebaut vorliegen, wieder abbauen. Damit wird das Schließen von n auf $n+1$ keineswegs schon benutzt — vielmehr ist es, wie wir schon von Dedekind her wissen und wie es meine Beweistheorie von neuem bestätigt, von hier noch ein weiter Weg und eine wesentlich andere Aufgabe, die Gültigkeit dieses Schlusses von n auf $n+1$ einzusehen.

Auf der Grundlage, die ich eben diskutiert habe, muß jedesmal der Beweis für die Widerspruchsfreiheit der Hinzunahme einer Aussage geführt werden. Gelingt es, diesen Beweis zu führen, so bedeutet dies für die Aussage, daß, wenn eine numerische und finit deutungsfähige Aussage aus ihr abgeleitet wird, diese tatsächlich jedesmal richtig ist. Der Beweis für die Widerspruchsfreiheit lehrt zugleich in jedem Falle eines zu einem falschen Resultat führenden Beweises, die Stelle zu finden, wo der Fehler liegt.

Es liegt auf der Hand, daß auch rein logische Probleme in den Bereich der eben von mir skizzierten Beweistheorie fallen. Als Beispiel diene folgendes Problem.

Problem IV.

Die Behauptung der Vollständigkeit des Axiomensystems der Zahlentheorie läßt sich auch so aussprechen: Wird eine der Zahlentheorie angehörige, aber nicht beweisbare Formel zu den Axiomen der Zahlentheorie hinzugenommen, so kann aus dem erweiterten Axiomensystem ein Widerspruch abgeleitet werden.

Da wir es hier in der Beweistheorie stets mit formalisierten Beweisen zu tun haben, so ist in den ausgesprochenen Behauptungen über die Vollständigkeit der Zahlentheorie zugleich die Behauptung eingeschlossen, daß die formalisierten Regeln des logischen Schließens jedenfalls im Gebiete der Zahlentheorie ausreichend sind.

Die Frage nach der Vollständigkeit des Systems der logischen Regeln, in allgemeiner Form gestellt, bildet ein Problem der theoretischen Logik. Zu dieser allgemeineren Problemstellung gelangen wir von der Zahlentheorie aus, wenn wir aus dem Bereich der betrachteten Formeln, und insbesondere auch der Axiome, alle diejenigen ausschalten, in denen das Zeichen „+1" vorkommt, dafür aber das Auftreten von Prädikaten-Variablen zulassen. Dies bedeutet sachlich, daß wir von dem Ordnungscharakter des Systems der Zahlen absehen und dieses nur als irgendein System von Dingen behandeln, auf welches Prädikate mit einem oder mehreren Subjekten bezogen werden können. Unter diesen Prädikaten wird nur die Gleichheit (Identität) als bestimmte Beziehung durch die üblichen Gleichheitsaxiome

$$a = a$$

$$a = b \rightarrow (A(a) \rightarrow A(b))$$

festgelegt, während die übrigen Prädikate frei wählbar bleiben.

In diesem Bereiche von Formeln sind diejenigen ausgezeichnet, die durch keine bestimmte Festlegung der wählbaren Prädikate widerlegbar sind. Diese stellen die allgemein gültigen logischen Sätze dar.

Es entsteht nun die Frage, ob alle diese Formeln aus den Regeln des logischen Schließens unter Hinzunahme der genannten Gleichheitsaxiome beweisbar sind, mit anderen Worten, ob das System der üblichen logischen Regeln vollständig ist.

Bisher haben wir durch Probieren die Überzeugung gewonnen, daß diese Regeln ausreichen. Ein wirklicher Nachweis dafür ist nur vorhanden im Bereich der reinen Aussagenlogik. Im Bereich der Logik der Prädikate mit einem Subjekt kann aus der Methode der Lösung des Entscheidungsproblems (Schrödersches Eliminationsproblem) ebenfalls ein Nachweis für die Vollständigkeit der Regeln gewonnen werden, wie dies in Anknüpfung an die Ansätze von Schröder zuerst von Löwenheim und später in abschließender Form von Behmann gezeigt worden ist.

Mein heutiger Vortrag zeigt, wie viele Probleme noch der Lösung harren. Aber im allgemeinen prinzipiellen Sinne ist auch nicht mehr die leiseste Spur einer Unklarheit möglich: jede prinzipielle Frage läßt sich auf Grund der von mir skizzierten Beweistheorie in einer Weise beantworten, die mathematisch präzise und eindeutig ist. Die betreffenden Sätze lassen sich zum Teil schon jetzt mit Hilfe der bisherigen Ergebnisse auf absolut sichere und rein mathematische Weise beweisen und sind daher dem Streite entzogen worden. Wer mich widerlegen will, muß, wie es in der Mathematik bisher stets üblich war und bleiben wird, mir genau die Stelle zeigen, wo der vermeintliche Fehler von mir liegt. Eine Einwendung, die das nicht tut, lehne ich a limine ab.

Ich glaube, daß meine Beweistheorie uns auch noch einen allgemeineren Dienst leistet. Denn wie wäre es mit der Wahrheit unseres Wissens überhaupt und wie mit der Existenz und dem Fortschritt der Wissenschaft bestellt, wenn es nicht einmal in der Mathematik sichere Wahrheit gäbe? Tatsächlich kommt heutzutage gar nicht selten selbst in Fachschriften und öffentlichen Vorträgen Zweifelsucht und Kleinmut gegenüber der Wissenschaft zum Ausdruck; es ist dies eine gewisse Art Okkultismus, den ich für schädlich halte. Die Beweistheorie macht eine solche Einstellung unmöglich und verschafft uns das Hochgefühl der Überzeugung, daß wenigstens dem mathematischen Verstande keine Schranken gezogen sind und daß er sogar die Gesetze des eigenen Denkens selbst aufzuspüren vermag. Cantor hat gesagt: das Wesen der Mathematik besteht in ihrer Freiheit, und ich möchte für die Zweifelsüchtigen und Kleinmütigen hinzufügen: in der Mathematik gibt es kein Ignorabimus, wir können vielmehr sinnvolle Fragen stets beantworten, und es bestätigt sich, was vielleicht schon Aristoteles vorausfühlte, daß unser Verstand keinerlei geheimnisvolle Künste treibt, vielmehr nur nach ganz bestimmten aufstellbaren Regeln verfährt, — zugleich die Gewähr für die absolute Objektivität seines Urteilens.

(Eingegangen am 25. 3. 1929.)

Unstarre geschlossene Flächen.

Von

Stefan Cohn-Vossen in Göttingen.

Eine Fläche heiße starr, wenn sie als Ganzes keine Infinitesimalverbiegungen außer infinitesimalen Bewegungen zuläßt; sonst unstarr. Die Eiflächen sind bekanntlich starr[1]). Aber auch unter den geschlossenen Flächen vom Geschlecht Null, deren Gaußsche Krümmung nicht überall positiv ist, war bisher keine unstarre bekannt (von einem trivialen Fall abgesehen)[2]). Im folgenden sollen nun solche unstarren Flächen angegeben werden, und zwar Rotationsflächen. Es zeigt sich, daß eine Kugel unstarr wird, wenn man längs eines Großkreises eine beliebig schmale und beliebig flache Rinne anbringt. Die Breite und Tiefe der Rinne sind aber nicht willkürlich, sondern müssen einem gewissen Wertevorrat entnommen sein, der eben unbegrenzt kleine Werte enthält.

Die Spezialisierung auf Rotationsflächen hat einen formalen Vorteil[3]): Für die Meridiane und Breitenkreise als Parameternetz reduziert sich das Starrheitsproblem durch Trennung der Veränderlichen auf *gewöhnliche* Differentialgleichungen, die den Verlauf der Verbiegung längs eines Meridians bestimmen (§ 1). Die Forderung, daß die Verbiegung auch in den Polen des Parameternetzes noch regulär sei (§ 2), führt — ähnlich wie in der Theorie der Kugelfunktionen — auf ein Sturm-Liouvillesches Eigenwertproblem, aber in einem anderen als dem üblichen Sinn. Die für unser Problem brauchbaren Eigenwerte sind nämlich von vornherein auf einen festen abzählbaren Vorrat beschränkt; der Meridian unserer Rotationsfläche muß so gewählt werden, daß das Sturm-Liouvillesche Problem mindestens

---

[1]) Vgl. W. Blaschke, Math. Zeitschr. 9 (1921) und Vorlesungen über Differentialgeometrie 1 (1921), Kap. V, § 77.

[2]) Jede Fläche, die ein ebenes Stück e enthält, ist unstarr; Biegungsvektor ist jeder Vektor, der für die Flächenpunkte außerhalb e verschwindet und für die von e senkrecht auf e steht. Die im folgenden betrachteten Infinitesimalverbiegungen verschwinden nie auf einem ganzen Flächenstück.

[3]) Vgl. Liebmann, Münchn. Berichte 1920.

einen Eigenwert aus jenem Vorrat besitzt; dann und nur dann ist die Fläche unstarr[4]).

Eine einfache Abschätzung mittels Sturmscher Sätze zeigt nun, daß zwar eine konvexe Kurve als Meridian nicht brauchbar sein kann (im Einklang mit der allgemeinen Theorie), daß man aber aus jeder konvexen Kurve durch Anbringen einer beliebig schwachen Ausbuchtung einen brauchbaren Meridian machen kann; je schwächer die Ausbuchtung, desto höher der zugehörige Eigenwert (§ 3).

In §§ 4, 5 werden unstarre Rotationsflächen von *Toruscharakter* aufgestellt. Das Parameternetz hat in diesem Falle keine Singularität. Dafür wird das Problem jetzt an den Breitenkreisen singulär, wo der Torus eine zur Achse senkrechte Tangentialebene hat. Beim Überschreiten solcher Stellen geht nämlich das Biegungsproblem vom elliptischen zum hyperbolischen Charakter über, und es zeigt sich, daß, grob ausgedrückt, nur „die Hälfte“ der im hyperbolischen Teil regulären Integrale heil über die Grenze kommt. Auf den übrigen parabolischen Breitenkreisen der Fläche (Wendepunkte der Meridiankurven) tritt diese Erscheinung nicht auf (§ 4).

Wir beschränken uns auf solche Torusflächen, die eine Symmetrieebene senkrecht zur Achse und nur zwei zur Achse senkrechte Tangentialebenen haben. Durch die Berührungspunkte mit diesen beiden Ebenen wird der Meridian in zwei Halbmeridiane zerlegt. Die Bestimmung einer Infinitesimalverbiegung, die auf einem solchen Halbmeridian bis in die Endpunkte hinein stetig ist, führt wieder auf ein Sturm-Liouvillesches Problem der oben beschriebenen Art.

Für beide von uns betrachteten Flächentypen ergibt sich (§§ 3, 5): Jede Fläche dieser Art ist Häufungsfläche einer Folge stetig gekrümmter *unstarrer* Flächen[5]). Die Punkte und auch noch die Tangentialebenen der Folge konvergieren gleichmäßig, die Krümmungen bleiben absolut beschränkt.

Eine *unstarre* Fläche braucht nicht *verbiegbar* zu sein. Das Problem der *Verbiegungen höherer Ordnung*[6]) wird in § 6 erörtert. Die unstarren Flächen der von uns betrachteten Art gestatten im allgemeinen auch eine Verbiegung zweiter Ordnung.

In § 7 werden zwei unstarre Flächen angegeben, die sich zu Modellzwecken eignen. Eine von ihnen besitzt als Meridian einen Streckenzug, besteht also nur aus Kegeln und Kegelstümpfen. Verbiegungen solcher geknickter Flächen sind meines Wissens noch nicht behandelt worden.

---

[4]) Auch Liebmann, loc. cit. [2]), stößt bei verwandten Fragen auf Eigenwertprobleme, aber im gewöhnlichen Sinn. Der Meridian ist vorgegeben.

[5]) Bei Deutung der Meridiankurven als Punkte eines Funktionenraums: Die Punkte, die unstarren Flächen entsprechen, liegen überall dicht.

[6]) Vgl. Liebmann, loc. cit. [2]) und Darboux, Théorie des surfaces 4, S. 2 bis 5.

§ 1.

Ansatz des Problems im Kleinen.

$\mathfrak{x}(u, v)$ sei der Ortsvektor der gesuchten unstarren Rotationsfläche, $\mathfrak{z}(u, v)$ der zugehörige Verschiebungsvektor (Bezeichnung wie bei Blaschke loc. cit. [1])).

Wir zerlegen $\mathfrak{x}$ und $\mathfrak{z}$ nicht nach einem festen Koordinatensystem in Komponenten, sondern nach einem beweglichen Dreikant, das dem Wesen der Rotationsflächen besser gerecht wird.

$\mathfrak{e}$ sei der Einheitsvektor parallel zur Drehachse, $\mathfrak{a} = \mathfrak{a}(v)$ ein auf $\mathfrak{e}$ senkrechter Einheitsvektor, der bei veränderlichem v einen Kreis mit $\mathfrak{e}$ als Achse und v als Bogenlänge beschreibt, d. h. der mit $\mathfrak{e}$ zusammen die Meridianebene der Fläche aufspannt. Also $\mathfrak{a}^2 = 1$, $\mathfrak{e}\mathfrak{a} = 0$, $\mathfrak{a}(v + 2\pi) = \mathfrak{a}(v)$. Dann gilt für $\mathfrak{a}' = \frac{d\mathfrak{a}}{dv}$: $\mathfrak{a}\mathfrak{a}' = \mathfrak{e}\mathfrak{a}' = 0$, $\mathfrak{a}'^2 = 1$, $\mathfrak{a}'' = -\mathfrak{a}$. $\mathfrak{e}, \mathfrak{a}, \mathfrak{a}'$ bilden unser Dreikant. Die Fläche $\mathfrak{x}(u, v)$ erhält dann die Gestalt:

$$\mathfrak{x}(u, v) = u\mathfrak{e} + r(u)\mathfrak{a}(v),$$

$u =$ konst. sind also die Breitenkreise, $v =$ konst. die Meridiane. Die Gestalt der Meridiane wird durch die Funktion $r(u)$ bestimmt; u ist nämlich der Abstand eines Flächenpunkts von einer festen Ebene senkrecht zur Achse, $r(u)$ ist der Abstand dieses Punkts von der Achse. Wir beschränken uns auf stetig gekrümmte Meridiane; dann sind r', r'' stetig außer in den Punkten, wo der Meridian eine zur Achse senkrechte Tangente hat $\left(\frac{1}{r'(u)} = 0\right)$. Diese Punkte schließen wir vorläufig aus und beschränken uns auf eine Zone $u_0 \leqq u \leqq u_1$ der Fläche, wo r, r', r'' stetig sind.

Der Verschiebungsvektor $\mathfrak{z}(u, v)$ habe nach $\mathfrak{e}, \mathfrak{a}, \mathfrak{a}'$ die Komponenten $\alpha(u, v)$; $\beta(u, v)$; $\gamma(u, v)$; also $\mathfrak{z} = \alpha\mathfrak{e} + \beta\mathfrak{a} + \gamma\mathfrak{a}'$. $\mathfrak{z}(u, v)$ wird bekanntlich durch die Gleichungen bestimmt:

$$\mathfrak{x}_u\mathfrak{z}_u = 0, \qquad \mathfrak{x}_v\mathfrak{z}_v = 0, \qquad \mathfrak{x}_u\mathfrak{z}_v + \mathfrak{x}_v\mathfrak{z}_u = 0. \tag{1}$$

Es ist

$$\begin{aligned} \mathfrak{x}_u &= \mathfrak{e} + r'\mathfrak{a}, & \mathfrak{x}_v &= r\mathfrak{a}', \\ \mathfrak{z}_u &= \alpha_u\mathfrak{e} + \beta_u\mathfrak{a} + \gamma_u\mathfrak{a}', & \mathfrak{z}_v &= \alpha_v\mathfrak{e} + (\beta_v - \gamma)\mathfrak{a} + (\beta + \gamma_v)\mathfrak{a}'. \end{aligned}$$

Das System (1) geht also über in

$$\begin{aligned} &\alpha_u + r'(u)\beta_u = 0, \\ &\beta + \gamma_v = 0, \\ &\alpha_v + r'(u)(\beta_v - \gamma) + r(u)\gamma_u = 0. \end{aligned} \tag{2}$$

Dies ist für α, β, γ ein lineares System, *dessen Koeffizienten von v unabhängig sind.* α, β, γ müssen von v periodisch mit der Periode 2π abhängen. Aus beiden Gründen setzen wir α, β, γ als Fourier-Reihen in v an:

$$\alpha(u,v) = \sum_{k=-\infty}^{+\infty} e^{ikv} \varphi_k(u); \quad \beta(u,v) = \sum_{k=-\infty}^{+\infty} e^{ikv} \psi_k(u); \quad \gamma(u,v) = \sum_{k=-\infty}^{+\infty} e^{ikv} \chi_k(u).$$

$\varphi_k, \psi_k, \chi_k$ sind komplexwertige Funktionen von u, die, damit α, β, γ reell werden, die Bedingung zu erfüllen haben:

$$\varphi_{-k} = \bar{\varphi}_k, \quad \psi_{-k} = \bar{\psi}_k, \quad \chi_{-k} = \bar{\chi}_k \quad (-\infty < k < +\infty).$$

Zunächst fragen wir nach Fourier-*Polynomen* α, β, γ, die (2) lösen. Die linken Seiten von (2) werden Fourier-Polynome vom selben Grad wie α, β, γ. Sämtliche Koeffizienten müssen verschwinden. Dies gibt für $\varphi_k, \psi_k, \chi_k$ das nur noch von u und dem ganzzahligen Parameter k abhängende System:

$$(3) \quad \begin{aligned} &\text{a)}\ \varphi_k'(u) + r'(u)\psi_k'(u) = 0, \\ &\text{b)}\ i k \chi_k(u) + \psi_k(u) = 0, \\ &\text{c)}\ i k \varphi_k(u) + r'(u)[i k \psi_k(u) - \chi_k(u)] + r(u)\chi_k'(u) = 0. \end{aligned}$$

Wir differenzieren c), eliminieren φ_k, ψ_k aus a), b) und erhalten dadurch:

$$(4) \qquad \boxed{r\chi_k'' + (k^2 - 1) r'' \chi_k = 0.}$$

Jedes nicht identisch verschwindende Integral $\chi_k^{(u)}$ von (4) für $k \geq 0$ liefert mittels (3) b), c) einen reellen nicht identisch verschwindenden Verschiebungsvektor [7])

$$\mathfrak{z}(u,v) = \mathfrak{z}_k(u,v)$$
$$= e^{ikv}(\varphi_k(u)\mathfrak{e} + \psi_k(u)\mathfrak{a} + \chi_k(u)\mathfrak{a}') + e^{-ikv}(\bar{\varphi}_k\mathfrak{e} + \bar{\psi}_k\mathfrak{a} + \bar{\chi}_k\mathfrak{a}').$$

Wegen der Linearität des Problems ist jede Summe $a_0\mathfrak{z}_0 + a_1\mathfrak{z}_1 + \dots + a_n\mathfrak{z}_n$ mit konstanten $a_0, \dots, a_n$ wieder ein Verschiebungsvektor.

Zunächst berechnen wir $\mathfrak{z}_0$. Aus (3) a) b) folgt für $k = 0$: $\psi_0 = 0$; $\varphi_0 = c_0 = \text{konst.}$; also nach (3) c):

$$r\chi_0' - r'\chi_0 = 0; \quad \chi_0 = C_0 r \qquad (C_0 = \text{konst.}),$$
$$\mathfrak{z}_0 = c_0\mathfrak{e} + C_0 r\mathfrak{a}'.$$

$\mathfrak{z}_0$ bedeutet ersichtlich eine infinitesimale Schraubung von beliebiger Ganghöhe $\frac{c_0}{C_0}$ um die Flächenachse (bzw. im Grenzfall $C_0 = 0$ eine Translation längs der Achse).

---

[7]) Die oben aufgestellte Realitätsforderung $\varphi_{-k} = \bar{\varphi}_k$ usw. wird hier benutzt. Sie erfüllt sich von selbst, denn (3) geht in sich über, wenn man φ usw. durch $\bar{\varphi}$ usw. und k durch $-k$ ersetzt.

Im Falle $k = 1$ erhalten wir ebenfalls eine infinitesimale Bewegung. $\mathfrak{z}_1$ durchläuft nämlich, wie der Leser leicht bestätigen kann, ein Stück einer durch $\mathfrak{e}$ gehenden Ebene; das bedeutet eine Schraubung bzw. Translation senkrecht zu dieser Ebene[8]). Daraus folgt umgekehrt, daß sich *jede* infinitesimale Bewegung als Linearform $a_0 \mathfrak{z}_0 + a_1 \mathfrak{z}_1$ schreiben läßt. Da $\mathfrak{z}_l$ und $\mathfrak{z}_k$ für $l \neq k$ linear unabhängig sind, kann $\mathfrak{z}_k$ für $k \geqq 2$ keine infinitesimale Bewegung bedeuten.

Jedes nicht identisch verschwindende Integral von (4) *für irgendein ganzzahliges* $k \geqq 2$ *liefert eine nichttriviale Infinitesimalverbiegung der betrachteten Zone unserer Fläche.*

§ 2.

Die Pole der Rotationsfläche.

Wir betrachten einen „Pol" der Rotationsfläche, d. h. einen ihrer Schnittpunkte mit der Achse. Die Fläche ist dort dann und nur dann stetig gekrümmt, wenn der Meridian dort auf der Achse senkrecht steht und stetig gekrümmt ist. Wir transformieren (3), (4) auf r als unabhängige Veränderliche, wobei wir $\varphi_k(r)$ usw. statt $\varphi_k[u(r)]$ usw. schreiben Dann geht (3) über in

$$
(5)\qquad
\begin{aligned}
&\text{a)}\quad \dot u(r)\,\dot\varphi_k(r) + \dot\psi_k(r) = 0\\
&\text{b)}\quad i k \chi_k(r) \quad + \psi_k(r) = 0\\
&\text{c)}\quad i k \dot u(r)\,\varphi_k(r) + i k \psi_k(r) - \chi_k(r) + r \dot\chi_k(r) = 0
\end{aligned}
$$

und aus (4) wird nach einfacher Zwischenrechnung

$$
(6)\qquad r \dot u(r)\, \ddot\chi_k(r) - r \ddot u(r)\, \dot\chi_k(r) - (k^2 - 1)\, \ddot u(r)\, \chi_k(r) = 0.
$$

Wir brauchen ein für $|r| \leqq r_0$ reguläres Lösungssystem von (5). Dabei müssen wir $\dot u(0) = 0$ voraussetzen, weil der Meridian die Achse senkrecht trifft. $\dot u(r)$ und $\ddot u(r)$ seien für $|r| \leqq r_0$ stetig. Wir setzen außerdem $\ddot u(r) \neq 0$ für $|r| \leqq r_0$ voraus, beschränken uns also auf Rotationsflächen, die in der Umgebung eines Pols keine parabolischen Punkte haben.

Setzen wir $\chi_k(r) = r^{\varrho_k} + r^{\varrho_k + 1} P(r)$, wo $P(r)$ für $|r| \leqq r_0$ zweimal stetig differenzierbar sein soll, so liefert (6) nach der Fuchsschen Theorie[9]) wegen $\ddot u(0) \neq 0$ die Gleichung

$$
\varrho_k(\varrho_k - 1) - \varrho_k - k^2 + 1 = 0,
$$

also

$$
\varrho_k = 1 \pm k.
$$

---

[8]) Vgl. Darboux, Théorie des surfaces 4, S. 9.

[9]) Vgl. Hoheisel, Gewöhnliche Differentialgleichungen (Samml. Göschen), S. 101 ff.

Jedes in $|r| \leqq r_0$ reguläre Grundintegral $\chi_k(r)$ von (6) verschwindet demnach in $r = 0$ von der Ordnung $k+1$. Die zugehörigen ψ_k, φ_k verschwinden dort nach (5) b) bzw. (5) c) von der Ordnung $k+1$ bzw. k. Der zugehörige Vektor $\mathfrak{z}_k(r, v)$ hat demnach wegen $k \geqq 2$ die Gestalt $\mathfrak{z}_k = r^2 \mathfrak{y}(r, v)$, wo der Vektor $\mathfrak{y}$ für $|r| \leqq r_0$ zweimal stetig differenzierbar nach r, v ist. Ersetzen wir also die Parameter r, v durch die Parameter $x = r\cos v$, $y = r \sin v$, deren Netz für $|r| \leqq r_0$ keine Singularität mehr hat, so ist $\mathfrak{z}_k$ nicht nur nach r, v, sondern auch nach x, y zweimal stetig differenzierbar.

Das[10]) *in* $r = 0$ *reguläre Grundintegral von* (6) *liefert für* $k \geqq 2$ *eine in der Umgebung des Pols zweimal stetig differenzierbare Infinitesimalverbiegung unserer Fläche.*

§ 3.

Aufstellung und Lösung des Sturm-Liouvilleschen Problems, das die unstarren Flächen liefert.

Wir gehen auf (3), (4) mit u als unabhängigen Veränderlichen zurück und wollen dies System längs eines ganzen Meridians integrieren. Seien die Pole der gesuchten Fläche etwa $u = \pm 1$, dann setzen wir von der Funktion $r(u)$ voraus, daß sie die Gestalt hat

$$r(u) = \sqrt{1 - u^2}\, l(u), \tag{7}$$

wo $l(u)$ eine im abgeschlossenen Intervall $-1 \leqq u \leqq +1$ positive, zweimal stetig differenzierbare Funktion ist; wir beschränken uns damit auf Flächen, deren Tangentialebene außer in den Polen nirgends auf der Achse senkrecht steht.

Aus §§ 1, 2 folgt:

Unsere Fläche ist unstarr, wenn (4) für irgendein ganzzahliges $k \geqq 2$ ein im ganzen Intervall $-1 \leqq u \leqq +1$ reguläres nicht identisch verschwindendes Integral besitzt.

Bei gegebener Funktion $r(u)$ *der Form* (7) *und verfügbarem nicht notwendig ganzzahligem* k *ist die Auffindung eines solchen Integrals von* (4) *und zugehöriger Eigenwerte* $k^2 - 1$ *ein Sturm-Liouvillesches Problem.* Wir haben umgekehrt die Aufgabe, $r(u)$ *so zu bestimmen, daß das Sturm-Liouvillesche Problem einen Eigenwert* $k^2 - 1$ *mit ganzzahligem* $k \geqq 2$ *besitzt.*

---

[10]) Jedes Grundintegral in einem singulären Punkt ist an sich nur bis auf einen konstanten Faktor bestimmt. Wir meinen hier und im folgenden dasjenige Grundintegral, in dem die niedrigste Potenz der unabhängigen Veränderlichen den Koeffizienten $+1$ hat.

Sei $r(u)$ irgendeine Funktion der Form (7), $\chi_k(u)$ hier und im folgenden das in $u = +1$ reguläre Grundintegral von (4). Dann geht $\chi_k(u)$ für $u \to -1$ im allgemeinen ins Unendliche. Wir zeigen nun: *Durch eine beliebig schwache bis zur 2. Ableitung stetige Abänderung von $r(u)$ in einem beliebig kleinen abgeschlossenen Teilintervall von $(-1, +1)$ kann man erreichen, daß gleichzeitig $\chi_k(u)$ (bei passend gewähltem, festem ganzzahligen $k \geqq 2$) in eine bei -1 reguläre Funktion übergeht. Die zu $r(u)$ gehörige Fläche wird also durch die Abänderung unstarr gemacht.*

Das Abänderungsverfahren verläuft am einfachsten bei konvexen Flächen (vgl. Fig. 1). Wir setzen also vorläufig voraus: $r'' < 0$ für $-1 < u < +1$. Dann gilt nach (4):

$$(8) \qquad \frac{\chi_k''(u)}{\chi_k(u)} > 0 \quad \text{für} \quad \chi_k(u) \neq 0 \qquad (-1 < u < +1).$$

Hieraus folgt, daß $\chi_k(u)$ kein positives Maximum und kein negatives Minimum in $(-1, +1)$ hat. Also kann $\chi_k(u)$ in $u = -1$ nicht verschwinden, also muß es dort unendlich werden[11]). (Das bestätigt den Satz von der Starrheit der Eiflächen für unseren Sonderfall.) Sei nun a irgendeine Zahl im Intervall $-1 < a < +1$. δ, ε seien beliebig klein vorgegebene positive Zahlen, insbesondere sei δ so klein, daß

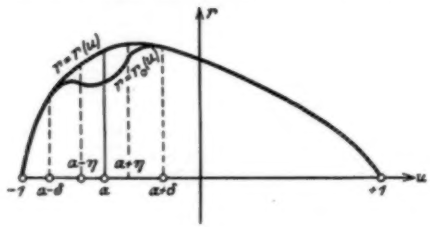


Fig. 1.

$$-1 < a - \delta < a + \delta < +1.$$

$r_0(u)$ sei eine Funktion, die folgende Bedingungen erfüllt:

1. $r_0(u) - r(u)$ ist zweimal stetig differenzierbar in $(-1, +1)$.
2. $r_0(u) - r(u) = 0$ außer in $(a - \delta,\ a + \delta)$.
3. $\underset{|u| \leqq 1}{\operatorname{Max}} |r_0(u) - r(u)| \leqq \varepsilon, \qquad \underset{|u| \leqq 1}{\operatorname{Max}} |r_0'(u) - r'(u)| \leqq \varepsilon.$
4. $\dfrac{r_0''(a)}{r_0(a)} = +1.$

Offenbar ist eine solche Funktion $r_0(u)$ konstruierbar. Zu ihr bestimmen wir ein η, $0 < \eta < \delta$, so daß

$$\frac{r_0''(u)}{r_0(u)} \geqq \frac{1}{2} \quad \text{für} \quad a - \eta \leqq u \leqq a + \eta.$$

[11]) Denn das zweite Grundintegral in $r = 0$ gehört zu dem negativen Exponenten $\varrho_2 = 1 - k$.

Sei $\chi_{k,0}(u)$ das für $a+\delta \leqq u \leqq +1$ mit $\chi_k(u)$ übereinstimmende Integral der Gleichung

$$(4_0) \qquad r_0 \chi''_{k,0} + (k^2-1) r''_0 \chi_{k,0} = 0,$$

die aus (4) entsteht, indem man r durch r_0 ersetzt.

Dann ist in $(a-\eta, a+\eta)$:

$$(9) \qquad \frac{\chi''_{k,0}(u)}{\chi_{k,0}(u)} \leqq -\frac{k^2-1}{2} \quad (\text{für } \chi_{k,0} \neq 0).$$

Die Funktion $y(u) = \sin\sqrt{\frac{k^2-1}{2}}\,u$ befriedigt die Gleichung

$$\frac{y''}{y} = -\frac{k^2-1}{2}.$$

Nach einem bekannten Sturmschen Satz [9]) ist der Abstand zweier konsekutiver Nullstellen von $\chi_{k,0}(u)$ in $(a-\eta, a+\eta)$ wegen (9) nicht größer als bei der Funktion $\sin\sqrt{\frac{k^2-1}{2}}\,u$. *Ist also k so groß gewählt, daß* $\sqrt{\frac{k^2-1}{2}} > \frac{\pi}{\eta}$, *so hat* $\chi_{k,0}(u)$ *mindestens eine Nullstelle in* $(a-\eta, a+\eta)$. k sei so gewählt und im folgenden fest.

Nunmehr betrachten wir die Funktion $r(u;t) = r_0(u) + t\,[r(u) - r_0(u)]$; wir machen also die Abänderung von $r(u)$ allmählich wieder rückgängig. $\chi(u;t)$ sei dasjenige Integral der Gleichung

$$(4_t) \qquad r(u;t)\chi''(u;t) + (k^2-1)\,r''(u;t)\,\chi(u;t) = 0$$

$\left(\frac{\partial F(u;t)}{\partial u} = F' \text{ gesetzt}\right)$, das für $a+\delta \leqq u \leqq +1$ mit $\chi_k(u)$ übereinstimmt, also $\chi(u;1) = \chi_k(u)$ wegen $r(u;1) = r(u)$. Dann hängt $\chi(u;t)$ für $a+\delta \geqq u > -1$ analytisch von t ab. Für hinreichend kleine t hat $\chi(u;t)$ mindestens eine Nullstelle in $(a-\eta, a+\eta)$. Also gibt es für kleine t eine stetige Funktion $b(t)$, so daß (vgl. Fig. 2):

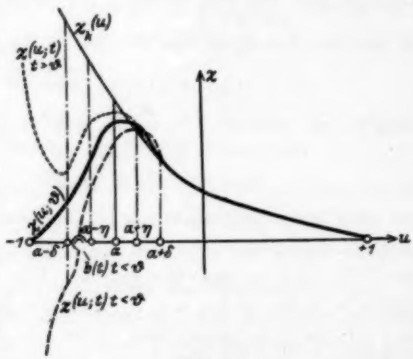


Fig. 2.

1. $-1 < b(t) < a+\delta$.
2. $\chi(u;t) = 0$ für $u = b(t)$.
3. $\chi(u;t) > 0$ für $b(t) < u < a+\delta$.

ϑ sei die obere Grenze aller solcher Zahlen τ $(\tau > 0)$, daß $b(t)$ für alle t im Intervall $0 \leqq t \leqq \tau$ existiert.

Dann ist $\vartheta \leqq +1$. Denn sonst hätte $\chi(u;1) = \chi_k(u)$ eine Nullstelle $b(1)$, $-1 < b(1) < a + \delta$. Das ist wegen (8) nicht der Fall.

Es ist $\lim\limits_{t \to \vartheta - 0} b(t) = -1$. Denn sonst wäre $\limsup\limits_{t \to \vartheta - 0} b(t) = b > -1$ und $b \leqq a + \delta$. Dann hätte $\chi(u;t)$ eine Nullstelle in beliebiger Nähe von b für $t > \vartheta$ und genügend kleines $t - \vartheta$.[12]) $b(t)$ wäre also für $0 \leqq t \leqq \tau$ mit $\tau > \vartheta$ definiert, im Widerspruch zur Maximaleigenschaft von ϑ.

$r(u;\vartheta)$ *ist die gesuchte Meridianfunktion, das Integral* $\chi(u;\vartheta)$ *von* (4_ϑ) *ist für* $u = -1$ *regulär.*

$\mathfrak{x}(u,v) = u\mathfrak{c} + r(u;\vartheta)\mathfrak{a}$ *ist eine unstarre Rotationsfläche.*

Zum Beweise genügt es zu zeigen, daß $\chi(u;\vartheta)$ für $-1 < u \leqq a + \delta$ *gleichmäßig in* u beschränkt ist[11]). Sei (für $0 \leqq t < \vartheta$) $M(t)$ das Maximum von $\chi(u;t)$ für festes t und für $b(t) \leqq u \leqq a + \delta$. *Dann bleibt* $M(t)$ *für* $t \to \vartheta$ *beschränkt.* $\chi(u;t)$ kann nämlich dieses Maximum $M(t)$ nur in $(a - \delta, a + \delta)$ annehmen, weil außerhalb dieses Intervalls überall (8) gilt und $\chi(b(t);t)$ verschwindet. Also ist in der Tat

$$0 < M(t) = \underset{\substack{b(t) \leqq u \leqq a+\delta \\ (0 \leqq t < \vartheta)}}{\operatorname{Max}} \chi(u;t) \leqq \underset{\substack{a-\delta \leqq u \leqq a+\delta \\ 0 \leqq t \leqq 1}}{\operatorname{Max}} \chi(u;t) = M.$$

Nun ist aber $\lim\limits_{t \to \vartheta - 0} b(t) = -1$. Für jeden Punkt c, $-1 < c < a + \delta$ gibt es also ein $\sigma > 0$, so daß für $\vartheta - \sigma < t < \vartheta$:

$$-1 < b(t) < c, \quad \text{also} \quad 0 < \chi(c;t) \leqq M.$$

Folglich ist auch $\chi(c;\vartheta) = \lim\limits_{t \to \vartheta - 0} \chi(c;t) \leqq M$, d. h.

$$0 < \chi(u;\vartheta) < M \quad \text{für} \quad -1 < u \leqq a + \delta. \qquad \text{Q. e. d.}$$

Wir lassen jetzt die Voraussetzung $r'' < 0$ fallen. Dann ist es nicht mehr möglich, analog dem Früheren den Punkt $u = a$ so zu wählen, daß $\chi_k(u) \neq 0$ für $-1 < u \leqq a$ *und für alle* k. Wir bestimmen diesmal $a < +1$ so, daß $r'' < 0$ für $-1 < u \leqq a$. Daß das geht, folgt aus (7). Jetzt gilt (8) in $-1 < u \leqq a$, $\chi_k(u)$ hat daher in diesem Intervall kein absolutes Extremum, also höchstens eine Nullstelle. Wir definieren $\delta, \varepsilon, r_0(u), \eta, \chi_{k,0}(u)$ wie im früheren. δ soll überdies so klein sein, daß $r'' < 0$ in $a \leqq u \leqq a + \delta$. Diesmal wählen wir k so groß, daß $\chi_{k,0}(u)$ mindestens *zwei* Nullstellen in $(a - \eta, a + \eta)$ hat. $r(u;t)$ und $\chi(u;t)$ seien wie oben erklärt. Wir wählen d so, daß $a + \delta < a + d < 1$, $\chi(u;t) = \chi_k(u) \neq 0$ für $u = a + d$, $r''(u) < 0$ für $a + \delta \leqq u \leqq a + d$. Dann gibt es

---

[12]) Bei dieser Schlußweise wird der Satz benutzt, daß ein nicht identisch verschwindendes Integral einer linearen homogenen Differentialgleichung zweiter Ordnung in einem regulären Punkt keine mehrfachen Nullstellen hat.

für kleine t zwei stetige Funktionen $b(t)$ und $B(t)$ mit den Eigenschaften:

1. $-1 < b(t) < B(t) < a + d$.
2. $\chi(u;t) = 0$ für $u = b(t)$ und für $u = B(t)$.
3. $\chi(u;t) \neq 0$ für $b(t) < u < B(t)$.

ϑ sei die obere Grenze aller solcher τ, daß $b(t)$ und $B(t)$ für $0 \leqq t \leqq \tau$ definiert sind. *Dann ist* $\vartheta \leqq 1$. Denn sonst hätte $\chi(u;1) = \chi_k(u)$ zwei Nullstellen in $(-1, a+d)$; das geht nicht, weil $\chi_k(u)$ für $-1<u<a+d$ kein absolutes Extremum hat. *Es ist* $B(t) \geqq a - \delta$, weil (8) für $\chi(u;t)$ in $-1 < u < a - \delta$ gilt, also $\chi(u;t)$ nicht zwei Nullstellen $b(t)$, $B(t)$ in diesem Intervall hat. *Es ist* $\lim\limits_{t \to \vartheta - 0} b(t) = -1$. Denn wäre $\limsup\limits_{t \to \vartheta - 0} b(t) = b > -1$, so könnten jedenfalls $b(t)$ und $B(t)$ für $t \to \vartheta - 0$ nicht zusammenrücken, weil $\chi(u;\vartheta)$ keine doppelte Nullstelle in b hat[12]). $B(t)$ könnte auch nicht nach $a+d$ rücken, denn es war $\chi(a+d;t) = \chi_k(a+d) \neq 0$ vorausgesetzt. Daher wären $b(t)$ und $B(t)$ über $t = \vartheta$ hinaus definierbar, entgegen der Definition von ϑ. $r(u;\vartheta)$ *ist die gesuchte Funktion*. Die gleichmäßige Beschränktheit von $\chi(u;\vartheta)$ folgt wie oben. Erstens sind die Funktionen $\chi(u;t)$ für $b(t) < u < B(t)$ gleichmäßig in t beschränkt, weil sie ihr Extrem nur in $(a-\delta, a+\delta)$ annehmen können; zweitens liegt jede Zahl c $(-1 < c < a - \delta)$ für $t \to \vartheta - 0$ zwischen $b(t)$ und $B(t)$, weil $\lim\limits_{t \to \vartheta - 0} b(t) = -1$, und $B(t) \geqq a - \delta$.

§ 4.

Parabolische Breitenkreise mit fester Stützebene.

Die parabolischen Breitenkreise einer Rotationsfläche zerfallen in drei Klassen. Erstens kann die Normalkrümmung in der Meridianrichtung verschwinden und in der Breitenkreisrichtung nicht; $r''(u) = 0$; $\frac{1}{r'(u)} \neq 0$ in der Bezeichnung von § 1. Zweitens kann die Normalkrümmung des Meridians von Null verschieden sein und die des Breitenkreises verschwinden; $\ddot{u}(r) \neq 0$, $\dot{u}(r) = 0$, $r \neq 0$ gemäß § 2. Drittens können beide Hauptkrümmungen zugleich verschwinden $(\ddot{u}(r) = \dot{u}(r) = 0)$. Der erste Fall spielt für die Integration von (4) keine Sonderrolle, denn $r''(u)$ ist nicht Koeffizient der höchsten Ableitung von χ_k. Im zweiten Fall dagegen, in dem wir (6) statt (4) zur Bestimmung von χ_k brauchen, hat (6) eine Singularität, denn $\dot{u}(r)$ ist Koeffizient von $\ddot{\chi}_k$. Der dritte Fall stellt eine noch höhere Singularität von (6) dar und soll unerörtert bleiben. Der zweite Fall trifft auf jeder torusförmigen Rotationsfläche mindestens für zwei Breitenkreise ein; nämlich für den höchsten und tiefsten, wenn die Achse vertikal steht.

Die Fläche hat längs eines solchen Breitenkreises eine und dieselbe Tangentialebene, die für die Umgebung des Kreises Stützebene ist.

Sei $r = 1$ ein solcher Kreis; $\dot{u}(1) = 0$, $\ddot{u}(1) \neq 0$. Wir wenden wie in § 2 die Fuchssche Theorie[9]) auf (6) im Punkte $r = 1$ an und setzen $\chi_k(r) = (r-1)^{\sigma_k} + (r-1)^{\sigma_k + 1} P(r-1)$. Dabei sei $P(x)$ eine für absolut kleines x zweimal stetig differenzierbare Funktion. Wegen $\ddot{u}(1) \neq 0$ erhält man aus (6) für σ_k die Gleichung

$$\sigma_k(\sigma_k - 1) - \sigma_k = 0; \qquad \sigma_k = \begin{cases} 2 \\ 0 \end{cases}.$$

Beide Grundintegrale des Punktes $r = 1$ sind also in dessen Umgebung beschränkt. Das zu $\sigma_k = 0$ gehörige scheidet für unseren Zweck aus. Denn es würde in $r = 1$ nicht verschwinden; also wäre nach (5) b) auch $\psi_k \neq 0$, dagegen nach (5) a) b) $\dot{\psi}_k = \dot{\chi}_k = 0$. Das ist mit (5) b) c) nicht verträglich.

Brauchbar in $r = 1$ ist also nur das Integral[10]), das dort wie $(r-1)^2$ verschwindet. Es liefert in der Tat mittels (5) b) c) eine in $r = 1$ reguläre Infinitesimalverbiegung der Fläche.

Hier sei eine Abschweifung gestattet. Fragen wir einmal nicht nach einer *speziellen* Verbiegung im *Großen*, sondern nach der *allgemeinsten* im *Kleinen* (Cauchy-Problem). $\mathfrak{z}$ und $\mathfrak{z}_u$ seien auf einem Breitenkreisbogen $u = c$; $v_1 \leqq v \leqq v_2$ so vorgegeben, daß (1) dort erfüllt ist. (1) soll über ein Gebiet $|u - c| \leqq c_0$, $v_1 \leqq v \leqq v_2$ integriert werden, das keinen Pol und keinen parabolischen Kreis zweiter Art enthalten möge. Nehmen wir an, $\mathfrak{z}$ und $\mathfrak{z}_u$ sind als Fourier-Polynome n-ten Grades auf $u = c$ vorgegeben. Dann kommt die Aufgabe darauf hinaus, (4) bzw. (6) für $k = 0, 1, 2, \ldots, n$ mit vorgegebenen $\chi_k(c)$, $\chi_k'(c)$ im Intervall $|u - c| \leqq c_0$ zu integrieren. Man erhält dann einen Biegungsvektor (vgl. § 1)

$$\mathfrak{z} = \mathfrak{z}_0 + \mathfrak{z}_1 + \cdots + \mathfrak{z}_n.$$

Für $n \to \infty$ wird die Entwicklung $\mathfrak{z} = \mathfrak{z}_0 + \mathfrak{z}_1 + \ldots + \mathfrak{z}_n + \ldots$ im allgemeinen überall außer in $u = c$ *divergent*, falls das Problem an der betrachteten Stelle elliptischen Charakter hat $\left(\frac{r''}{r} < 0\right)$. Sie *konvergiert* aber für $|u - c| \leqq c_0$, wenn die Komponenten von $\mathfrak{z}$, $\mathfrak{z}_u$ *analytische Funktionen von v mit der Periode* 2π sind. Man zeigt mittels Sturmscher Sätze, daß c_0 nur von den vorgegebenen Funktionen und von $\liminf \frac{r''}{r}$ abhängt[13]).

[13]) Die $\chi_k(u)$ haben für $\frac{r''}{r} < 0$ die Größenordnung $a_k e^{\left|\frac{r''}{r}\right| \sqrt{k^2 - 1}\, u}$, wie man aus (4) entnimmt. Dabei sind die a_k durch die Vorgaben in $u = c$ bestimmt. Konvergenz wird also dann und nur dann eintreten, wenn die $|a_k|$ wie ϑ^k $(0 < \vartheta < 1)$ abnehmen. Das ist aber gerade die notwendige und hinreichende Bedingung dafür, daß $\sum_{-\infty}^{+\infty} a_k e^{ikv}$ eine *analytische* Funktion von v ist.

Insbesondere endet die Konvergenz im allgemeinen nicht auf parabolischen Kreisen erster Art; man kann das Cauchy-Problem mit analytischen Vorgaben vom elliptischen ins hyperbolische Gebiet hineinintegrieren.

Enthält aber das gewünschte Integrationsgebiet einen parabolischen Kreis zweiter Art, so ist das Cauchy-Problem nicht mehr allgemein lösbar, auch wenn man sich mit Fourier-Polynomen als Vorgaben begnügt. Zwischen den Vorgaben $\mathfrak{z}$, $\mathfrak{z}_u$ muß eine Relation bestehen, damit das zugehörige Integral von (6) zu den im parabolischen Kreis regulären gehört. Für $n \to \infty$ bleibt diese Einschränkung a fortiori bestehen. Die Konvergenzfrage liegt wie im ersten Fall.

Für die lineare homogene *partielle* Differentialgleichung der Infinitesimalverbiegung vollzieht sich also in unserem (separierbaren) Falle der Übergang vom elliptischen ins hyperbolische Gebiet in verschiedener Weise, je nachdem die parabolische Kurve vom ersten oder vom zweiten Typ ist. Dieses verschiedene Verhalten dürfte kaum an der Rotationssymmetrie unseres Problems liegen. Ich kenne aber keine analoge Unterscheidung in allgemeineren Fällen.

§ 5.

Unstarre Torusflächen.

Wir beschränken uns auf solche Torusflächen, die nur zwei zur Achse senkrechte Tangentialebenen haben; diese berühren den Torus längs zweier parabolischer Kreise zweiter Art. Außerdem möge der Torus eine Symmetrieebene senkrecht zur Achse haben. Diese sei $u = 0$, die parabolischen Kreise seien $u = \pm 1$, $r = a$ $(a > 0)$. Dann hat der Meridian die Gestalt

$$(10)\qquad \begin{cases} r = r(u) \;\;= \;\;\sqrt{1-u^2}\, l(u) + a, \\ r = r^*(u) = -\sqrt{1-u^2}\, l^*(u) + a. \end{cases}$$

Dabei seien $l(u)$, $l^*(u)$ für $-1 \leqq u \leqq +1$ zweimal stetig differenzierbar. Ferner sei $r(u) > 0$, $r^*(u) > 0$ für $-1 \leqq u \leqq +1$, $l(-u) = l(u)$, $l^*(-u) = l^*(u)$, $l(1) = l^*(1) > 0$.

Jede solche Fläche kann durch eine beliebig kleine, die Symmetrie erhaltende Änderung unstarr gemacht werden.

Seien nämlich $\chi_k(u)$, $\chi_k^*(u)$ die zu $r(u)$ bzw. $r^*(u)$ gehörigen, in $u = +1$ regulären Grundintegrale von (4). Dann soll die Abänderung von $r(u)$ und von $r^*(u)$ im Intervall $+1 > u \geqq 0$ so erfolgen, daß in diesem Intervall je eine Nullstelle von $\chi_k(u)$ und von $\chi_k^*(u)$ in den Punkt $u = 0$ rücken. Im Intervall $0 \geqq u > -1$ werde die zur ersten spiegelbildliche Änderung des Meridians vorgenommen. Dann folgt $\chi_k(-u) = -\chi_k(u)$, $\chi_k^*(-u) = -\chi_k^*(u)$; also werden insbesondere $\chi_k(u)$,

$\chi_k^*(u)$ regulär in $u = -1$. Damit ist unsere Aufgabe gelöst. Die Konstruktion des Biegungsvektors erfolgt gemäß (3) wie in § 3.

Das Änderungsverfahren ist analog dem von § 3, modifiziert durch unsere Symmetrieannahme; als Abänderungszone wählen wir gerade die Umgebung von $u = 0$. Wir betrachten zunächst den Halbmeridian $r = r(u)$. Wegen $r(-u) = r(u)$ ist $r'(0) = 0$. Zur Abkürzung setzen wir $r(0) = b$, $\left|\frac{r''(0)}{r(0)}\right| = c$. Nun sei ε, $0 < \varepsilon < 1$, beliebig vorgegeben. Dann sei δ so bestimmt, daß $0 < \delta < \frac{\varepsilon}{4(b+1)(c+1)}$ und daß für $0 \leqq u \leqq \delta$:
$|r(u) - b| \leqq \frac{\varepsilon}{2}$, $|r'(u)| \leqq \frac{\varepsilon}{3}$, $\left|\frac{r''(u)}{r(u)}\right| < c + 1$. Nun bestimmen wir $r_0(u)$ so, daß

1. $r_0(u) - r(u)$ zweimal stetig differenzierbar für $0 \leqq u \leqq 1$.
2. $r_0(u) - r(u) = 0$ für $\delta \leqq u \leqq 1$.
3. $\underset{0 \leqq u \leqq \delta}{\operatorname{Max}} |r_0(u) - b| \leqq \frac{\varepsilon}{2}$, $\underset{0 \leqq u \leqq \delta}{\operatorname{Max}} |r_0'(u)| \leqq \frac{2\varepsilon}{3}$

(also $\underset{0 \leqq u \leqq \delta}{\operatorname{Max}} |r_0(u) - r(u)| \leqq \varepsilon$, $\underset{0 \leqq u \leqq \delta}{\operatorname{Max}} |r_0'(u) - r'(u)| \leqq \varepsilon$).

4. $r_0(u) = b + 2(c+1)(b+1)u^2$ für $0 \leqq u \leqq \frac{\delta}{2}$.

4. ist mit 3. verträglich. Denn es ist für $0 \leqq u \leqq \frac{\delta}{2}$ nach 4.:

$$\operatorname{Max} |r_0'| = 2(b+1)(c+1)\delta < \frac{\varepsilon}{2},$$

also nach dem Mittelwertsatz $\underset{0 \leqq u \leqq \frac{\delta}{2}}{\operatorname{Max}} |r_0 - b| < \frac{\delta}{2} \cdot \frac{\varepsilon}{2} < \frac{\varepsilon}{2}$. Im Intervall $\frac{\delta}{2} \leqq u \leqq \delta$ kann man $r_0(u)$ offenbar so erklären, daß auch dort 1., 2., 3. gelten.

Ferner gilt

$$\underset{0 \leqq u \leqq \frac{\delta}{2}}{\operatorname{Min}} \frac{r_0''(u)}{r_0(u)} > 4 \underset{0 \leqq u \leqq \delta}{\operatorname{Max}} \frac{r''(u)}{r(u)}. \tag{11}$$

In der Tat ist nach unseren Festsetzungen $M = \underset{0 \leqq u \leqq \delta}{\operatorname{Max}} \frac{r''(u)}{r(u)} < c + 1$, $m = \underset{0 \leqq u \leqq \frac{\delta}{2}}{\operatorname{Min}} \frac{r_0''(u)}{r_0(u)} \geqq \frac{4(b+1)(c+1)}{b+\varepsilon}$, also $m > 4M$. Setzen wir $\frac{1}{2}\sqrt{m} - \sqrt{M} = \sigma$, so ist $\sigma > 0$.

Nun sei $\chi_{k,0}(u)$ wie in § 3 erklärt als das Integral von (4_0) § 3, das für $\delta \leqq u \leqq 1$ mit $\chi_k(u)$ übereinstimmt. Wie in § 3 schließt man: $\chi_{k,0}(u)$ hat im Intervall $0 \leqq u \leqq \frac{\delta}{2}$ mindestens $n = \frac{\delta}{2\pi}\sqrt{m(k^2-1)}$ Nullstellen. Dagegen hat $\chi_k(u)$ in dem größeren Intervall $0 \leqq u \leqq \delta$ höchstens $N = 1 + \frac{\delta}{\pi}\sqrt{M(k^2-1)}$ Nullstellen. Es ist $n - N = \frac{\delta}{\pi}\sigma\sqrt{k^2-1} - 1$.

Ist k groß genug, so ist $n - N \geqq 1$; dann hat $\chi_{k,0}(u)$ im Intervall $0 \leqq u \leqq \delta$ mindestens eine Nullstelle mehr als $\chi_k(u)$. k sei so groß gewählt und im folgenden fest. Wie in § 3 definieren wir nun $r(u;t)$ und $\chi(u;t)$: $r(u;t) = r_0(u) + t[r(u) - r_0(u)]$, $\chi(u;t)$ das Integral von (4_t) § 3, das für $\delta \leqq u \leqq 1$ mit $\chi_k(u)$ übereinstimmt.

Dann gibt es ein ϑ, $0 \leqq \vartheta \leqq 1$, *so daß* $\chi(0;\vartheta) = 0$. Wäre nämlich $\chi(u;t) \neq 0$ für alle t im Intervall $0 \leqq t \leqq 1$, so könnte die Anzahl n_t der Nullstellen von $\chi(u;t)$ im Intervall $0 \leqq u \leqq \delta$ sich für $0 \leqq t \leqq 1$ nicht ändern. Denn diese wandern stetig mit t, mehrere fallen nie zusammen[11]), und $\chi(\delta;t)$ ist konstant gleich $\chi_k(\delta)$. Andererseits ist $\chi(u;0) = \chi_{k,0}(u)$, $\chi(u;1) = \chi_k(u)$, also $n_0 \geqq n$, $n_1 = N$, also $n_0 - n_1 \geqq 1$. Das ist ein Widerspruch, der die Existenz unseres ϑ beweist.

Erklären wir $r(u;\vartheta)$ für $0 \geqq u \geqq -1$ durch die Gleichung $r(-u;\vartheta) = r(u;\vartheta)$, so ist nach dem vorher Bemerkten die Kurve $r = r(u;\vartheta)$ der äußere Halbmeridian einer unstarren Torusfläche.

Den inneren Halbmeridian erzeugt man aus $r^*(u)$ auf genau dieselbe Art. —

Wenn man die Symmetrievoraussetzung fallen läßt, macht es zwar keine Schwierigkeit, beide Halbmeridiane so zu deformieren, daß sowohl $\chi_k(u)$ als auch $\chi_k^*(u)$ in $u = \pm 1$ regulär werden. Aber der Zusammenschluß beider Funktionen in $u = \pm 1$ genügt unseren Forderungen im allgemeinen nicht. Wir verlangen (Bezeichnung wie in § 2): $\ddot{\chi}_k(a) = \ddot{\chi}_k^*(a)$ für $u = +1$ und für $u = -1$.

Wenn man $\chi_k(u)$ gegebenenfalls durch konst. $\chi_k(u)$ ersetzt, kann man für $u = +1$ sicher jene Gleichung erzwingen; dann gilt sie aber im allgemeinen für $u = -1$ nicht, während bei unserem symmetrischen Ansatz die eine Hälfte die andere nach sich zieht. Der allgemeine Fall läßt sich vermutlich dadurch erledigen, daß man das Intervall, in dem man den einen Halbmeridian abändert, zunächst verschiebbar in $(-1, +1)$ annimmt, um es dann an eine passende Stelle zu rücken.

§ 6.

Verbiegungen höherer Ordnung.

Das Problem der Infinitesimalverbiegung $\mathfrak{z}(u,v)$ einer Fläche $\mathfrak{x}(u,v)$ ordnet sich bekanntlich[8]) dem folgenden allgemeineren Problem unter. Die Vektoren $\mathfrak{z}^{(1)}, \mathfrak{z}^{(2)}, \ldots, \mathfrak{z}^{(n)}$ sind so zu bestimmen, daß das Linienelement der Flächen $\mathfrak{x}(u,v;t) = \mathfrak{x}(u,v) + \sum_{l=1}^{n} t^l \mathfrak{z}^{(l)}(u,v)$ (das jedenfalls ein Polynom $2n$-ten Grades in t ist) keine Glieder mit $t, t^2, \ldots, t^n$ enthält. $\mathfrak{z}^{(l)}(u,v)$ heißt Biegungsvektor l-ter Ordnung von $\mathfrak{x}(u,v)$; der von uns

bisher betrachtete Vektor $\mathfrak{z}(u, v)$ bestimmt also die „Verbiegung erster Ordnung“; $\mathfrak{z} = \mathfrak{z}^{(1)}$. Die $\mathfrak{z}^{(n)}$ bestimmen sich aus $\mathfrak{z}^{(1)}$ durch die Rekursionsformeln

$$2\, d\mathfrak{x}\, d\mathfrak{z}^{(n)} = -\sum_{l=1}^{n-1} d\mathfrak{z}^{(l)}\, d\mathfrak{z}^{(n-l)},$$

d. h.

$$(12)\qquad \begin{gathered} 2\,\mathfrak{x}_u \mathfrak{z}_u^{(n)} = -\sum_{l=1}^{n-1} \mathfrak{z}_u^{(l)} \mathfrak{z}_u^{(n-l)}; \quad 2\,\mathfrak{x}_v \mathfrak{z}_v^{(n)} = -\sum_{l=1}^{n-1} \mathfrak{z}_v^{(l)} \mathfrak{z}_v^{(n-l)}; \\ \mathfrak{x}_u \mathfrak{z}_v^{(n)} + \mathfrak{x}_v \mathfrak{z}_u^{(n)} = -\sum_{l=1}^{n-1} \mathfrak{z}_u^{(l)} \mathfrak{z}_v^{(n-l)}. \end{gathered}$$

Gelingt es, diese Differentialgleichungen *für alle* n zu integrieren, und konvergiert $\sum_{n=1}^{\infty} t^n \mathfrak{z}^{(n)}$ für kleine $t \neq 0$, so ist

$$\mathfrak{x}(u, v; t) = \mathfrak{x}(u, v) + \sum_{n=1}^{\infty} t^n \mathfrak{z}^{(n)}(u, v)$$

eine in t analytische Schar isometrischer Flächen mit $\mathfrak{x}(u, v; 0) = \mathfrak{x}(u, v)$; $\mathfrak{x}(u, v)$ läßt eine stetige Verbiegung zu.

Kann man nach unserer Methode unstarre Rotationsflächen herstellen, die Verbiegungen höherer Ordnungen besitzen, oder die gar analytisch verbiegbar sind?

Die Rotationsfläche $\mathfrak{x}(u, v)$ vom Geschlecht der Kugel sei unstarr und ihr Biegungsvektor erster Ordnung sei nach § 1

$$\mathfrak{z}^{(1)}(u, v) = \mathfrak{z}(u, v) = \alpha \mathfrak{e} + \beta \mathfrak{a} + \gamma \mathfrak{a}'.$$

Setzen wir an:

$$\mathfrak{z}^{(2)} = \alpha^{(2)} \mathfrak{e} + \beta^{(2)} \mathfrak{a} + \gamma^{(2)} \mathfrak{a}',$$

so folgt aus (12):

$$\begin{aligned} \alpha_u^{(2)} + r' \beta_u^{(2)} &= -\tfrac{1}{2}\,[\alpha_u^2 + \beta_u^2 + \gamma_u^2], \\ \gamma_v^{(2)} + \beta^{(2)} &= -\tfrac{1}{2}\,[\alpha_v^2 + (\beta_v - \gamma)^2], \\ \alpha_v^{(2)} + r'(\beta_v^{(2)} - \gamma^{(2)}) + r\gamma_u^{(2)} &= -\tfrac{1}{2}\,[\alpha_u \alpha_v + \beta_u(\beta_v - \gamma)]. \end{aligned}$$

Die linken Seiten sind in $\alpha^{(2)}, \beta^{(2)}, \gamma^{(2)}$ so gebaut wie die von (2) in α, β, γ. Rechts stehen quadratische Ausdrücke in α, β, γ. Daraus folgt: Sind α, β, γ Fourier-Polynome in v vom Grade m, so sind die $\alpha^{(2)}, \beta^{(2)}, \gamma^{(2)}$ Fourier-Polynome in v vom Grade $2m$. Nehmen wir an, (4) sei nur für $k = m\,(\geqq 2)$ regulär auf der ganzen Fläche lösbar, für $k \neq m$, $k \geqq 2$ nicht. Dann haben α, β, γ die Form

$$\alpha = \varphi_m(u)\, e^{imv} + \bar{\varphi}_m(u)\, e^{-imv} \quad \text{usw.}$$

$\alpha^{(2)}, \beta^{(2)}, \gamma^{(2)}$ müssen also die Form haben:

$$\alpha^{(2)} = \varphi_{2m}^{(2)} e^{2imv} + \varphi_0^{(2)} + \bar{\varphi}_{2m}^{(2)} e^{-2imv} \quad \text{usw.}$$

Entsprechend für den zugehörigen Biegungsvektor dritter Ordnung

$$\mathfrak{z}^{(3)} = \alpha^{(3)}\mathfrak{e} + \beta^{(3)}\mathfrak{a} + \gamma^{(3)}\mathfrak{a}'.$$

$\alpha^{(3)}, \beta^{(3)}, \gamma^{(3)}$ müssen die Gestalt haben:

$$\alpha^{(3)} = \varphi_{3m}^{(3)}(u) e^{3imv} + \varphi_m^{(3)}(u) e^{imv} + \bar{\varphi}_m^{(3)}(u) e^{-imv} + \bar{\varphi}_{3m}^{(3)}(u) e^{-3imv} \quad \text{usw.}$$

Das Problem der Verbiegung zweiter Ordnung wird analog § 1 äquivalent mit der Integration der Differentialgleichung

$$r\chi_k^{(l)\prime\prime} + (k^2-1)\, r''\chi_k^{(l)} = R_k^{(l)}(u), \tag{13}$$

und zwar für $l = 2$, $k = 0$ und $k = 2m$. $R_k^{(2)}(u)$ ist ein längerer Ausdruck in $\varphi_k, \psi_k, \chi_k, \chi_k', \chi_k''$. Die genauere Betrachtung lehrt mittels der Fuchsschen Theorie, daß (13) für jedes $k \geqq 2$, $k \neq m$ sowie für $k = 0$ ein und nur ein überall reguläres Integral besitzt[14]). *Jede Rotationsfläche vom Geschlecht der Kugel, die eine Verbiegung erster Ordnung gestattet, aber keine zwei linear unabhängigen, gestattet auch eine Verbiegung zweiter Ordnung.*

Verbiegung *dritter* Ordnung ist äquivalent mit Integration von (13) für $l = 3$, $k = m$ und $k = 3m$. Die rechten Seiten sind jetzt bilinear in $\varphi, \psi, \chi, \varphi^{(2)}, \psi^{(2)}, \chi^{(2)}$ und deren Ableitungen. Für $k = 3m$ macht die Integration keine Schwierigkeit, wohl aber für $k = m$. Denn nach unserer Annahme hat das *homogene* Problem (4) für $k = m$ eine Lösung, also das *inhomogene* (13) nur, wenn die rechte Seite eine Integralbedingung erfüllt. Ich weiß nicht, was diese Integralbedingung für den Meridian bedeutet, ob sie z. B. identisch erfüllt oder vielleicht unerfüllbar ist.

Falls es eine Verbiegung dritter Ordnung gibt, so auch eine vierter. Denn diese erfordert die Integration von (13) nur für $k = 0$, $2m$, $4m$, nicht für $k = m$. Ebenso erfordert allgemein das Verbiegungsproblem $(2n+1)$-ter Ordnung eine Integralbedingung, dasjenige $2n$-ter nicht.

Andere Erscheinungen treten auf, wenn (4) für mehrere k lösbar ist, wenn also die Fläche mehrere linear unabhängige Verbiegungen erster Art zuläßt. Besonders wäre der Fall zu untersuchen, daß (4) für *alle* k lösbar ist, daß also die Fläche *als Ganzes* eine ebenso große Mannigfaltigkeit von Infinitesimalverbiegungen gestattet wie eine beliebig kleine Umgebung eines ihrer Pole.

---

[14]) Man muß dabei berücksichtigen, daß $R_k^{(l)}(u)$ in den Polen $r = 0$ hinreichend stark verschwindet.

§ 7.

Zwei Beispiele unstarrer Flächen.

Daß eine starre Fläche durch eine beliebig kleine Abänderung unstarr gemacht werden kann, erscheint zunächst überraschend. In Wirklichkeit bedeutet aber die Unstarrheit nicht viel. Unterwirft man nämlich die unstarre Fläche $\mathfrak{x}(u, v)$ mit dem Biegungsvektor $\mathfrak{z}$ der Transformation $\mathfrak{x}(u, v) \to \mathfrak{x}(u, v) + t\,\mathfrak{z}(u, v)$, und geht dabei ein Bogen s der Fläche in einen Bogen σ über, so folgt aus der Unstarrheit nur die Beziehung

$$s < \sigma < s\sqrt{1 + M t^2}.$$

Dabei ist $M = \operatorname{Max} \frac{d\mathfrak{z}^2}{d\mathfrak{x}^2}$. Nur für $|t| < \frac{1}{M}$ macht es sich fühlbar, daß bei unserer Transformation die Änderung der Bogenlänge von geringerer Größenordnung ist als die mittlere Gestaltsänderung der Fläche. Setzen wir $\operatorname{Max} |\mathfrak{x}(u, v) - \mathfrak{x}(u_1, v_1)| = A$, $\operatorname{Max} |\mathfrak{z}(u, v) - \mathfrak{z}(u_1, v_1)| = B$ (wo (u, v) und (u_1, v_1) beliebige Punkte der Fläche $\mathfrak{x}$ sind), so kann man den Ausdruck

$$U = \frac{B^2}{A^2 M}$$

als Maß der Unstarrheit der Fläche $\mathfrak{x}$ bezeichnen. Man überschlägt leicht aus (4) und (3), daß U für große k die Größenordnung $\frac{1}{k^2}$ hat. Nun gehören die Biegungsvektoren, die das Verfahren aus §§ 3, 5 liefert, im allgemeinen *zu sehr großen* k. Die dort erzeugten Flächen sind also nur sehr schwach unstarr. Bei Deformationen, die einen merklichen Bruchteil des Flächendurchmessers betragen, ändern sich die Bogenlängen einer solchen Fläche nicht merklich weniger als die einer starren Fläche.

Im folgenden sollen dagegen zwei unstarre Rotationsflächen vom Geschlecht der Kugel angegeben werden, deren Infinitesimalverbiegung *zu* $k = 2$ gehört. Man müßte ihre Unstarrheit an Modellen normaler Größe bemerken können[15]).

1. Beispiel (Fig. 3).

Es sei, mit zunächst unbestimmten Konstanten a, b, c:

1.	$r = \sqrt{u}$	für $0 \leq u \leq \frac{1}{4}$	; also	$\frac{r''}{r} = -\frac{1}{4u^2}$
2.	$r = \frac{1}{2} e^{2u - \frac{1}{2}}$	„ $\frac{1}{4} \leq u \leq a$	; „	$\frac{r''}{r} = 4$
3.	$r = b \cos(u - a - c)$	„ $a \leq u \leq a + c$;	„	$\frac{r''}{r} = -1.$

[15]) Durch die Transformation $r \to c_1 r$, $u \to c_2 u$ (c_1, c_2 beliebige nichtverschwindende Zahlen) kann man die Gestalt der Flächen noch affin modifizieren. Denn (4) ist gegen solche Transformationen invariant.

4. $r(u) = r(2a + 2c - u)$ „ $a + c \leqq u \leqq 2a + 2c$; die Fläche besitzt die Symmetrieebene $u = a + c$.

Die zweiten Ableitungen dieses Meridians haben Sprungstellen in $u = \frac{1}{4}$, $u = a$, $u = a + c$, $u = a + 2c$, $u = 2a + 2c - \frac{1}{4}$. Die in dieser Arbeit nur für stetig gekrümmte Meridiane angestellten Betrachtungen lassen sich auf Meridiane mit bloß stückweise stetiger Krümmung ohne weiteres übertragen, weil in der ursprünglichen Form (1) des Biegungsproblems keine höheren als erste Ableitungen miteinander verknüpft werden.

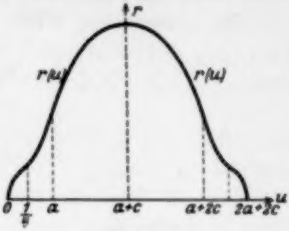


Fig. 3.

a, b, c sind so zu bestimmen, daß $r(u)$ und $r'(u)$ an den Nahtstellen stetig bleiben. Für $u = \frac{1}{4}$ und $u = 2a + 2c - \frac{1}{4}$ ist das von selbst der Fall, für $u = a$ erhält man

$$\frac{1}{2} e^{2a - \frac{1}{2}} = b \cos c; \qquad e^{2a - \frac{1}{2}} = b \sin c.$$

Also $\underline{\operatorname{tg} c = 2}$; $\underline{b = \frac{\sqrt{5}}{2} e^{2a - \frac{1}{2}}}$; a bleibt noch unbestimmt. Für $\chi_k(u)$ *und* $k = 2$ kann man nach (4) ansetzen (A, B, C zunächst unbestimmt):

$$\begin{array}{lll}
1.\ \chi_2(u) = u^{\frac{3}{2}} & \text{für} & 0 \leqq u \leqq \frac{1}{4} \\
2.\ \chi_2(u) = A \cos(2\sqrt{3}\, u + B) & „ & \frac{1}{4} \leqq u \leqq a \\
3.\ \chi_2(u) = C \operatorname{Cof}[\sqrt{3}(u - a - c)] & „ & a \leqq u \leqq a + c.
\end{array}$$

Damit $\chi_2(u)$, $\chi_2'(u)$ an den Nahtstellen stetig bleiben, muß gelten:

$$\frac{1}{8} = A \cos\left(\frac{\sqrt{3}}{2} + B\right); \quad \frac{3}{4} = -2\sqrt{3} A \sin\left(\frac{\sqrt{3}}{2} + B\right)$$

$$A \cos(2\sqrt{3}\, a + B) = C \operatorname{Cof}(\sqrt{3}\, c); \quad 2A \sin(2\sqrt{3}\, a + B) = C \operatorname{Sin}(\sqrt{3}\, c).$$

Die beiden ersten Gleichungen liefern A, B:

$$\underline{\operatorname{tg}\left(\frac{\sqrt{3}}{2} + B\right) = -\sqrt{3}. \ A = \frac{1}{4}.}$$

Die letzten beiden Gleichungen liefern nunmehr a als Funktion von c und B

$$\underline{\operatorname{tg}(2\sqrt{3}\, a + B) = \frac{1}{2} \operatorname{Tg} \sqrt{3}\, c.}$$

Endlich haben wir für C:

$$\underline{C = \frac{1}{2\sqrt{4 \operatorname{Cof}^2(\sqrt{3}\, c) + \operatorname{Sin}^2(\sqrt{3}\, c)}}.}$$

Der Ansatz war wie in § 5 auf Symmetrie angelegt. Für $u = a + c$ ist $\chi_2'(u) = r'(u) = 0$. Setzt man daher $\chi_2(u) = \chi_2(2a + 2c - u)$ für $a + c \leqq u \leqq 2a + 2c$, so ist $\chi_2(u)$ auf der ganzen Fläche zweimal stückweise stetig differenzierbar; die Fläche ist unstarr.

Die numerischen Werte der Konstanten sind:

$$a \sim 0{,}6806 \qquad A = \frac{1}{4}$$
$$b \sim 2{,}645 \qquad B \sim 1{,}229$$
$$c \sim 1{,}107 \qquad C \sim 0{,}0709.$$

2. Beispiel (Fig. 4).

Die Fläche besitzt wieder eine zur Achse senkrechte Symmetrieebene. Der Meridian ist aber diesmal ein Streckenzug; die Fläche besteht aus zwei Kreiskegeln und aus zwei Kreiskegelstümpfen, ist also besonders leicht durch ein Modell zu verwirklichen.

An den Knickstellen σ des Meridians versagt unsere bisherige Theorie. Hier tritt an Stelle von (4) die Eckenbedingung

$$(14) \qquad \chi_k'(u)\Big|_{\sigma-0}^{\sigma+0} = -(k^2-1)\frac{\chi_k(\sigma)}{r(\sigma)}\, r'(u)\Big|_{\sigma-0}^{\sigma+0}.$$

Man erhält (14) aus (3) c), indem man verlangt, daß $\varphi_k(u)$ an der Knickstelle stetig bleibt (man kann (14) aber auch durch Grenzübergang aus (4) ableiten). Gilt (14) und ist auf den geradlinigen Stücken gemäß (4) $\chi_k''(u) = 0$, so ist (3) auf der ganzen Fläche durch ein System stetiger stückweise linearer Funktionen erfüllt, liefert also einen Biegungsvektor $\mathfrak{z}$, der, außer höchstens in den Polen, stetig und stückweise stetig differenzierbar ist. An den Polen tritt an Stelle der Betrachtungen von § 2 die Forderung $\varphi_k(u) = \psi_k(u) = \chi_k(u) = 0$. Sie ist für die Stetigkeit von $\mathfrak{z}$ an den Polen notwendig und hinreichend, wie man ohne weiteres verifizieren kann.

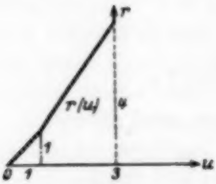


Fig. 4.

Nun setzen wir, bei vorläufig unbestimmtem a, b, und für $k = 2$:

$$
\begin{aligned}
r &= u && \text{für } 0 \leqq u \leqq 1\\
r &= \frac{3}{2}u - \frac{1}{2} && \text{„ } 1 \leqq u \leqq a\\
r(u) &= r(2a - u) && \text{„ } a \leqq u \leqq 2a\\
\chi_2(u) &= u && \text{„ } 0 \leqq u \leqq 1\\
\chi_2(u) &= bu + 1 - b && \text{„ } 1 \leqq u \leqq a.
\end{aligned}
$$

Zur Bestimmung von b hat man (14):

$$b-1=-3\left(\frac{3}{2}-1\right);\quad \underline{b=-\frac{1}{2}}.$$

Jetzt bestimmen wir a so, daß $\chi_2(a)=0$. Also $-\frac{a}{2}+\frac{3}{2}=0$; $\underline{a=3}$. Setzen wir $\chi_2(u)=-\chi_2(2a-u)$ für $a\leqq u\leqq 2a$, so erfüllt $\chi(u)$ (14) für $u=a$ und offenbar alle übrigen Forderungen für $0\leqq u\leqq 2a$. Unsere Fläche ist also in der Tat unstarr.

Einen unstarren Meridian liefert auch der Wert $a=\frac{7}{8}$. (14) ist nämlich in $u=a=\frac{7}{8}$ erfüllt, wenn man $\chi_2(u)=+\chi_2(2a-u)$ definiert. Für $u=a=\frac{7}{8}$ ist $\varphi_2(u)=0$, wie aus (3) c) zu verifizieren. Das bedeutet aber: Die $\mathfrak{c}$-Komponente von $\mathfrak{z}$ verschwindet in der Symmetrieebene; jede von dieser abgeschnittene Flächenhälfte gestattet also eine „Gleitverbiegung" nach Liebmann[8]).

Alle Komponenten von $\mathfrak{z}$ haben wegen $k=2$ die Periode π; es ist am Modell schön zu beobachten, wie die Fläche nachgibt, wenn man sie *an zwei diametralen Punkten* des größten Breitenkreises zusammendrückt, wie sie dagegen widersteht, wenn man an drei oder mehr äquidistanten Punkten dieses Kreises zu drücken sucht. Ebenso gestattet sie keine Relativbewegung ihrer Spitzen, weil dort $\varphi_k(u)=\psi_k(u)=\chi_k u=0$, also $\mathfrak{z}_k=0$ für $k\geqq 2$.

Modelle für $k=3$ müssen natürlich auf Druck in *drei* äquidistanten Punkten eines Breitenkreises ansprechen.

(Eingegangen am 15. 12. 1928.)

Zur Fortsetzbarkeit gewisser Dirichletscher Reihen.

Von

Kurt Mahler in Krefeld.

In der Arbeit: „Über analytische Funktionen und die Verteilung der Zahlen mod. Eins“ [1]) bewies E. Hecke für quadratische Irrationalzahlen o, daß die Dirichletsche Reihe

$$\sum_{n=1}^{\infty} \frac{no-[no]-\frac{1}{2}}{n^s} \qquad (s=\sigma+ti)$$

sich als meromorphe Funktion von s in die ganze Ebene fortsetzen läßt und allein an den Stellen

$$s=-2k+\frac{2k^*\pi i}{\log\eta} \qquad \begin{pmatrix} k=0,1,2,\ldots \\ k^*=0,\mp 1,\mp 2,\ldots \end{pmatrix}$$

Pole höchstens erster Ordnung haben kann; dabei bedeutet η eine gewisse positive Einheit des durch o erzeugten Zahlkörpers.

Einen ganz anderen Beweis für den gleichen Satz lieferten G. H. Hardy und J. E. Littlewood in der gemeinsamen Arbeit: „The analytic character of a Dirichlet's series considered by Hecke“ [2]).

Im folgenden wird ein weiterer Beweis dargestellt, und zwar für den allgemeineren Satz:

Seien $f(x_1 x_2)$ und $g(x_1 x_2)$ zwei Polynome mit reellen Koeffizienten vom Grad m bzw. n; $f^{(m)}(x_1 x_2)$ der höchste homogene Bestandteil von $f(x_1 x_2)$; $r(x_1 x_2)=f(x_1 x_2)-f^{(m)}(x_1 x_2)$; die natürliche Zahl d sei gleich 2, *wenn $r(x_1 x_2)$ identisch verschwindet, sonst gleich* 1. *Es gelte*

$$f(x_1 x_2)>0 \quad \textit{für} \quad x_1 \geqq 1,\ x_2 \geqq 1;$$

$$f^{(m)}(x_1 x_2)>0 \quad \textit{für} \quad x_1 \geqq 0,\ x_2 \geqq 0,\ x_1+x_2>0.$$

---

[1]) Abh. a. d. math. Sem. d. Hamb. Univ. **1** (1922), S. 54.

[2]) Abh. a. d. math. Sem. d. Hamb. Univ. **3** (1924), S. 57.

o sei eine positive quadratische Irrationalzahl, ζ die erzeugende positive Einheit des Zahlkörpers $K(o)$.

Die Funktion

$$Z_o(s \mid f(x_1 x_2),\ g(x_1 x_2)) = \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} g(x_1 x_2) f(x_1 x_2)^{-s}$$

läßt sich meromorph in die ganze Ebene fortsetzen; sie hat einfache Pole höchstens an den Stellen

$$s = \frac{n+2-k}{m} \qquad \begin{pmatrix} k = 0, 1, 2, \ldots \\ k^* = 0, \mp 1, \mp 2, \ldots \end{pmatrix}$$
$$s = \frac{n-dk}{m} + \frac{2k^* \pi i}{m \log \eta}$$

und zwar bedeutet η eine ganze Potenz von ζ. In jedem endlichen Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2$$

gilt gleichmäßig in s bei beliebig kleinem $\varepsilon > 0$:

$$Z(s \mid f(x_1 x_2),\ g(x_1 x_2)) = O(e^{\varepsilon |t|}),$$

wenn s um mindestens eine feste positive, beliebig kleine Strecke von den vorigen Polen entfernt bleibt. —

Die Hauptmittel des Beweises sind ein bekannter Satz von Mellin und der folgende Hilfssatz:

Es sei l eine natürliche Zahl, ω eine positive und ω' eine komplexe Konstante. Die Reihe

$$Z(s, l \mid \omega, \omega') = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \left(\frac{x_1 + \omega' x_2}{x_1 + \omega x_2}\right)^l (x_1 + \omega x_2)^{-s}$$

läßt sich meromorph in die ganze Ebene fortsetzen; sie hat einfache Pole höchstens an den Stellen

$$s = 2, 1, 0, \ldots, 2 - l.$$

In jedem endlichen Streifen

$$2 - k < \sigma_1 \leqq \sigma \leqq \sigma_2,$$

aus dem um jedem der Pole herum eine beliebig kleine Umgebung entfernt ist, besteht gleichmäßig in s und l die Ungleichung

$$|Z(s, l \mid \omega, \omega')| \leqq c \left(2 \frac{1 + |\omega'|}{\min(1, \omega)}\right)^l (1 + |t|)^{2k},$$

wo die Zahl $c > 0$ allein von $\omega, \omega', \sigma_1, \sigma_2$ abhängt.

Der Beweis dieses Satzes wird im ersten Kapitel dargestellt; man benutzt dabei die bekannte Eulersche Summenformel. Auch Hardy und Littlewood machen in ihrer Arbeit von einem ähnlichen Hilfssatz Gebrauch, der jedoch viel tiefer und schwerer zu beweisen ist.

Im zweiten Kapitel wird für einen beliebigen irrationalen echten Bruch

$$o = \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots, \qquad o = \frac{1}{a_1 + o_1}, \qquad o_1 = \frac{1}{a_2 + o_2}, \qquad o_2 = \frac{1}{a_3 + o_3}, \ldots$$

die Identität

$$Z_o(s \mid f(x_1 x_2), g(x_1 x_2)) = Z_\nu(s \mid f(x_1 x_2), g(x_1 x_2)) +$$
$$+ (-1)^\nu Z_{o_\nu}(s \mid f_\nu(x_1 x_2), g_\nu(x_1 x_2))$$

nachgewiesen; dabei bedeutet $Z_\nu(s \mid f(x_1 x_2), g(x_1 x_2))$ eine meromorphe Funktion mit Polen 1. Ordnung allein bei

$$s = \frac{n + 2 - k}{m} \qquad (k = 0, 1, \ldots),$$

es ist p_ν und q_ν der ν-te Näherungsnenner und -zähler von o und

$$z_1^{(\nu)} = p_\nu z_1 + q_\nu z_2, \qquad z_2^{(\nu)} = p_{\nu-1} z_1 + q_{\nu-1} z_2,$$
$$f_\nu(z_1 z_2) = f(z_1^{(\nu)}, z_2^{(\nu)}), \qquad g_\nu(z_1 z_2) = g(z_1^{(\nu)}, z_2^{(\nu)}).$$

Im dritten Kapitel wird o als quadratische Irrationalzahl mit reinperiodischem Kettenbruch angenommen: $o = o_\nu$ und ν gerade. Aus dem obigen Hilfssatz und der vorigen Funktionalgleichung folgt dann der Satz:

Die Reihe

$$Z_o(s, l) = \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} \left(\frac{x_1 + o' x_2}{x_1 + o x_2}\right)^l (x_1 + o x_2)^{-s},$$

wo o' die zu o konjugierte Irrationalzahl ist, läßt sich in die ganze Ebene meromorph fortsetzen, hat einfache Pole höchstens bei

$$s = 2, 1, 0, \ldots, 2 - l; \qquad s = -2l + \frac{2k^* \pi i}{\log \eta}, \qquad \eta = p_\nu + p_{\nu-1} o$$
$$(k^* = 0, \mp 1, \mp 2, \ldots)$$

und befriedigt gleichmäßig in jedem endlichen Streifen

$$2 - k < \sigma_1 \leqq \sigma \leqq \sigma_2 \qquad (k = 0, 1, 2, \ldots)$$

die Ungleichung

$$|Z_o(s, l)| \leqq \Gamma C^l (1 + |t|)^{2k}$$

mit zwei positiven Konstanten C und Γ, die von s und l nicht abhängen, wenn s von den Polen mindestens eine feste positive, beliebig kleine Strecke entfernt bleibt.

Im vierten Kapitel wird gleichfalls $o = o_\nu$ angenommen und

$$g(x_1 x_2) = (x_1 + o x_2)^{n-l} (x_1 + o' x_2)^l,$$

dagegen $f(x_1 x_2)$ gleich einer Form. Dann existiert zu einer beliebig großen natürlichen Zahl N eine Zahl Λ_{N} und eine Form m-ten Grades

$\mathfrak{F}_{\mathsf{N}}(x_1 x_2)$, so daß

$$f_{\mathsf{N}}(x_1 x_2) = \Lambda_{\mathsf{N}}((x_1 + o x_2)^m + F_{\mathsf{N}}(x_1 x_2))$$

ist, während mit wachsendem N alle Koeffizienten von $F_{\mathsf{N}}(x_1 x_2)$ gegen Null streben. Für großes $\mathsf{N} = N\nu$ ist somit

$$Z_0(s \mid f(x_1 x_2),\ g(x_1 x_2)) = \mathsf{Z}_{\mathsf{N}}(s \mid f(x_1 x_2),\ g(x_1 x_2)) +$$
$$+ \Lambda_{\mathsf{N}}^{-s}\, \eta^{N(n-2l)} \sum_{\lambda=0}^{\infty} \binom{-s}{\lambda} Z_0(m(s+\lambda) \mid x_1 + o x_1,\ g(x_1 x_2)\, \mathfrak{F}_{\mathsf{N}}(x_1 x_2)^{\lambda})$$

und aus dieser Entwicklung läßt sich nach dem Ergebnis des dritten Kapitels die Fortsetzbarkeit herleiten.

Im letzten Kapitel lassen sich durch einige triviale Schlüsse die letzten Beschränkungen über o, $f(x_1 x_2)$ und $g(x_1 x_2)$ aufheben und es folgt dann endlich der angegebene allgemeine Satz.

I.

1. Ein allgemeiner Satz von Mellin sagt aus[8]):

„Die beiden Polynome

$$f(x_1 \ldots x_p) = \sum_{h=0}^{m} f^{(h)}(x_1 \ldots x_p), \qquad g(x_1 \ldots x_p) = \sum_{k=0}^{n} g^{(k)}(x_1 \ldots x_p)$$

mit reellen Koeffizienten seien vom Grad m und n in den x; die Formen

$$f^{(h)}(x_1 \ldots x_p), \quad g^{(k)}(x_1 \ldots x_p)$$

bedeuten ihre homogenen Bestandteile der Dimension h und k. Es gelte

$$f(x_1 \ldots x_p) > 0 \quad \text{für} \quad x_1 \geqq 0, \ldots, x_p \geqq 0;$$

$$f^{(m)}(x_1 \ldots x_p) > 0 \quad \text{für} \quad x_1 \geqq 0, \ldots, x_p \geqq 0, \quad \sum_{l=1}^{p} x_l > 0.$$

Dann konvergiert die Summe

$$Z(s) = \sum_{x_1=1}^{\infty} \cdots \sum_{x_p=1}^{\infty} g(x_1 \ldots x_p)\, f(x_1 \ldots x_p)^{-s}, \qquad (s = \sigma + ti)$$

für

$$\sigma > \frac{n+p}{m}$$

und läßt sich in die ganze Ebene als meromorphe Funktion fortsetzen, mit einfachen Polen allein bei

$$s = \frac{n+p-k}{m} \qquad \begin{pmatrix} k = 0, 1, 2, \ldots \\ s \neq 0, -1, -2, \ldots \end{pmatrix}.$$

[8]) Hj. Mellin, Eine Formel für den Logarithmus transzendenter Funktionen von endlichem Geschlecht, Acta soc. sc. fenn. **29** (1900), Nr. 4. Ein neuer Beweis mit der obigen Größenabschätzung steht in meiner Arbeit: Über einen Satz von Mellin, Math. Annalen **100** (1928), S. 384.

In jedem endlichen Streifen

$$\frac{n+p-k}{m} < \sigma_1 \leqq \sigma \leqq \sigma_2 \qquad (k = 0, 1, 2, \ldots)$$

gilt gleichmäßig in σ für große $|t|$:

$$Z(s) = O(|t|^{pk})."$$

Genau die gleichen Eigenschaften hat auch die Summe

$$Z^*(s) = \sum_{x_1=0}^{\infty} \cdots \sum_{x_p=0}^{\infty} g(x_1 \ldots x_p)\, f(x_1 \ldots x_p)^{-s},$$

denn sie ergibt sich durch Addition endlichvieler Funktionen der vorigen Art. Folglich gilt das gleiche auch für die ursprüngliche Funktion

$$Z(s) = \sum_{x_1=1}^{\infty} \cdots \sum_{x_p=1}^{\infty} g(x_1 \ldots x_p)\, f(x_1 \ldots x_p)^{-s},$$

wenn nur

$$f(x_1 \ldots x_p) > 0 \quad \text{für} \quad x_1 \geqq 1, \ldots, x_p \geqq 1;$$

$$f^{(m)}(x_1 \ldots x_p) > 0 \quad \text{für} \quad x_1 \geqq 0, \ldots, x_p \geqq 0, \quad \sum_{l=1}^{p} x_l > 0$$

angenommen wird.

2. Es ist nötig, noch einige Eigenschaften der speziellen Mellinschen Funktion

$$Z(s, l \mid \omega_1 \omega_2) = Z(s, l) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} \left(\frac{x_1 + \omega' x_2}{x_1 + \omega x_2}\right)^l (x_1 + \omega x_2)^{-s}$$

herzuleiten; dabei sei ω eine positive, ω' eine beliebige Konstante und l eine nichtnegative ganze rationale Zahl. Die Eulersche Summenformel liefert die Umformung (k = bel. nat. Zahl)[4]):

$$Z(s, l) = \left(\frac{1+\omega'}{1+\omega}\right)^l (1+\omega)^{-s} + \sum_{l_1=0}^{k} \sum_{l_2=0}^{k} J_{l_1 l_2}^{(k)}(s, l),$$

$$J_{l_1 l_2}^{(k)}(s, l) = \iint_{\mathfrak{S}} H_{l_1 l_2}^{(k)}(x_1 x_2) \frac{\partial^{l_1+l_2}}{\partial x_1^{l_1} \partial x_2^{l_2}} \left\{\left(\frac{x_1 + \omega' x_2}{x_1 + \omega x_2}\right)^l (x_1 + \omega x_2)^{-s}\right\} dx_1\, dx_2.$$

Die Punktmenge $\mathfrak{S}$ ist hier durch

$$x_1 \geqq 0, \quad x_2 \geqq 0, \quad \max(x_1 x_2) \geqq 1$$

definiert; die $(k+1)^2$ Funktionen

$$H_{l_1 l_2}^{(k)}(x_1 x_2) \qquad \begin{pmatrix} l_1 = 0, 1, 2, \ldots, k \\ l_2 = 0, 1, 2, \ldots, k \end{pmatrix}$$

---

[4]) Vergleiche zu der hier verwendeten Form der Eulerschen Summenformel meine Arbeit [5]). Mit den dortigen Bezeichnungen ist hier gleichzusetzen:

$$H_{l_1 l_2}^{(k)}(x_1 x_2) = h_{l_1}^{(k)}(x_1) \cdot h_{l_2}^{(k)}(x_2).$$

sind von s, l, ω, ω' unabhängig, in allen Punkten von $\mathfrak{S}$ absolut genommen unterhalb einer allein von k abhängigen Konstanten, und auch noch von x_1 und x_2 unabhängig, wenn

$$l_1 < k, \quad l_2 < k$$

ist.

3. Es sei r eine positive Veränderliche, φ eine Veränderliche im Intervall

$$0 \leqq \varphi \leqq \frac{\pi}{2}.$$

Mit $\xi_1(\varphi)$, $\xi_2(\varphi)$, $\varrho(\varphi)$ seien folgende drei stetigen Funktionen bezeichnet:

$$\xi_1(\varphi) = \begin{cases} 1\,, \\ \operatorname{ctg}\varphi\,, \end{cases} \quad \xi_2(\varphi) = \begin{cases} \operatorname{tg}\varphi\,, \\ 1\,, \end{cases} \quad \varrho(\varphi) = \begin{cases} \dfrac{1}{\cos^2\varphi} \\ \dfrac{1}{\sin^2\varphi} \end{cases} \text{für} \quad \begin{matrix} 0 \leqq \varphi \leqq \dfrac{\pi}{4}, \\ \dfrac{\pi}{4} \leqq \varphi \leqq \dfrac{\pi}{2}. \end{matrix}$$

Wenn

$$x_1 = r\,\xi_1(\varphi), \quad x_2 = r\,\xi_2(\varphi)$$

gesetzt wird, so ist die Funktionaldeterminante dieser Transformation gleich

$$\frac{d(x_1 x_2)}{d(r\varphi)} = \varrho(\varphi)\cdot r.$$

Die Menge aller Punkte (x_1, x_2) mit

$$r \geqq 1, \quad 0 \leqq \varphi \leqq \frac{\pi}{2}$$

ist mit $\mathfrak{S}$ identisch.

Man zeigt leicht die Gleichung

$$\frac{\partial^{l_1+l_2}}{\partial x_1^{l_1}\partial x_2^{l_2}}\left\{\left(\frac{x_1+\omega' x_2}{x_1+\omega x_2}\right)^l (x_1+\omega x_2)^{-s}\right\} =$$

$$= (l_1+l_2)!\sum_{\lambda_1=0}^{l_1}\sum_{\lambda_2=0}^{l_2}\frac{\binom{l_1}{\lambda_1}\binom{l_2}{\lambda_2}}{\binom{l_1+l_2}{\lambda_1+\lambda_2}}\,\omega^{l_2-\lambda_2}\,\omega'^{\lambda_2}\binom{l}{\lambda_1+\lambda_2}\binom{-s}{l_1+l_2-\lambda_1-\lambda_2}\left(\frac{x_1+\omega' x_2}{x_1+\omega x_2}\right)^{l-\lambda_1-\lambda_2}(x_1+\omega x_2)^{-s-l}$$

Aus ihr ergibt sich die Zerlegung

$$J^{(k)}_{l_1 l_2}(s, l) =$$

$$= (l_1+l_2)!\sum_{\lambda_1=0}^{l_1}\sum_{\lambda_2=0}^{l_2}\frac{\binom{l_1}{\lambda_1}\binom{l_2}{\lambda_2}}{\binom{l_1+l_2}{\lambda_1+\lambda_2}}\,\omega^{l_2-\lambda_2}\,\omega'^{\lambda_2}\binom{l}{\lambda_1+\lambda_2}\binom{-s}{l_1+l_2-\lambda_1-\lambda_2} J^{*(k)}_{l_1 l_2}(s+l_1+l_2,\ l-\lambda_1-\lambda_2)$$

$$J^{*(k)}_{l_1 l_2}(s, l) = \iint\limits_{\mathfrak{S}} H^{(k)}_{l_1 l_2}(x_1 x_2)\left(\frac{x_1+\omega' x_2}{x_1+\omega x_2}\right)^l (x_1+\omega x_2)^{-s}\,dx_1\,dx_2.$$

4. Es sei erstens

$$0 \leqq l_1 \leqq k-1, \quad 0 \leqq l_2 \leqq k-1.$$

Da $H^{(k)}_{l_1 l_2}(x_1 x_2)$ konstant ist, so kann in diesem Fall die Integration nach r ausgeführt werden, so daß sich der Ausdruck ergibt:

$$(2-s-l_1-l_2)\,J^{(k)}_{l_1 l_2}(s,l) =$$

$$= (l_1+l_2)!\,H^{(k)}_{l_1 l_2}(x_1 x_2)\sum_{\lambda_1=0}^{l_1}\sum_{\lambda_2=0}^{l_2}\frac{\binom{l_1}{\lambda_1}\binom{l_2}{\lambda_2}}{\binom{l_1+l_2}{\lambda_1+\lambda_2}}\,\omega^{l_1-\lambda_1}\omega'^{\lambda_2}\binom{l}{\lambda_1+\lambda_2}\binom{-s}{l_1+l_2-\lambda_1-\lambda_2}\times$$

$$\times J^*(s+l_1+l_2,\; l-\lambda_1-\lambda_2),$$

$$J^*(s,l) = \int_0^{\frac{\pi}{2}} \varrho\left(\frac{\xi_1+\omega'\xi_2}{\xi_1+\omega\xi_2}\right)^l(\xi_1+\omega\xi_2)^{-s}\,d\varphi.$$

Ihm entnimmt man:

Die Funktion

$$J^{(k)}_{l_1 l_2}(s,l) = \iint_{\mathfrak{S}} H^{(k)}_{l_1 l_2}(x_1 x_2)\,\frac{\partial^{l_1+l_2}}{\partial x_1^{l_1}\,\partial x_2^{l_2}}\left\{\left(\frac{x_1+\omega' x_2}{x_1+\omega x_2}\right)^l(x_1+\omega x_2)^{-s}\right\}dx_1\,dx_2$$

$$(l_1, l_2 = 0, 1, \ldots, k-1)$$

ist meromorph und hat einen einzigen einfachen Pol höchstens an der Stelle

$$s = 2 - l_1 - l_2, \qquad \text{falls } l_1 + l_2 \leqq l.$$

In jedem endlichen Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2,$$

aus dem eine beliebig kleine Umgebung des Poles entfernt ist, besteht die Ungleichung

$$|J^{(k)}_{l_1 l_2}(s,l)| \leqq c_1\cdot\left(2\cdot\frac{1+|\omega'|}{\min(1,\omega)}\right)^l(1+|t|)^{l_1+l_2-1},$$

und die Konstante c_1 hängt hier allein von ω, ω', k, σ_1, σ_2 ab.

Zweitens sei mindestens eine der Zahlen l_1 und l_2 gleich k. In diesem Fall konvergiert das Integral

$$J^{*(k)}_{l_1 l_2}(s,l) = \int_1^{\infty} dr \int_0^{\frac{\pi}{2}} H^{(k)}_{l_1 l_2}(r\xi_1, r\xi_2)\left(\frac{\xi_1+\omega'\xi_2}{\xi_1+\omega\xi_2}\right)^l(\xi_1+\omega\xi_2)^{-s}r^{1-s}\varrho\,d\varphi$$

gewiß absolut und gleichmäßig in jeder Halbebene

$$2 < \sigma_1 \leqq \sigma$$

und stellt folglich hier eine reguläre Funktion dar. Aus der Darstellung von $J^{(k)}_{l_1 l_2}(s,l)$ als Summe solcher Integrale folgt, daß in jedem Streifen

$$2 - l_1 - l_2 < \sigma_1 \leqq \sigma \leqq \sigma_2$$

die Ungleichung

$$|J^{(k)}_{l_1 l_2}(s,l)| \leqq c_2\left(2\,\frac{1+|\omega'|}{\min(1,\omega)}\right)^l (1+|t|)^{l_1+l_2}$$

besteht, wobei die Konstante c_2 allein von ω, ω', k, σ_1, σ_2 abhängt.

5. Die beiden vorigen Ergebnisse wenden wir auf die Formel

$$Z(s,l)=\left(\frac{1+\omega'}{1+\omega}\right)^l (1+\omega)^{-s}+\sum_{l_1=0}^{k}\sum_{l_2=0}^{k} J^{(k)}_{l_1 l_2}(s,l)$$

an; man gelangt so zu dem Resultat:

Die Reihe

$$Z(s,l\mid\omega_1\omega_2)=Z(s,l)=\sum_{x_1=1}^{\infty}\sum_{x_2=1}^{\infty}\left(\frac{x_1+\omega' x_2}{x_1+\omega x_2}\right)^l (x_1+\omega x_2)^{-s}$$

läßt sich als meromorphe Funktion in die ganze Ebene fortsetzen; sie hat einfache Pole höchstens an den Stellen

$$s=2,1,0,-1,-2,\ldots,2-l.$$

In jedem endlichen Streifen

$$2-k<\sigma_1\leqq\sigma\leqq\sigma_2 \qquad (k=0,1,2,\ldots),$$

aus dem um jedem der Pole eine beliebig kleine Umgebung entfernt ist, besteht die Ungleichung

$$|Z(s,l)|\leqq c\left(2\,\frac{1+|\omega'|}{\min(1,\omega)}\right)^l (1+|t|)^{2k}$$

mit einer Zahl c, die allein von ω, ω', k, σ_1, σ_2 abhängt.

II.

6. Es seien

$$f(x_1x_2)=\sum_{h=0}^{m} f^{(h)}(x_1x_2),\qquad g(x_1x_2)=\sum_{k=0}^{n} g^{(k)}(x_1x_2)$$

zwei Polynome mit reellen Koeffizienten vom Grad m und n in den x; die Formen $f^{(h)}(x_1x_2)$ und $g^{(k)}(x_1x_2)$ bedeuten ihre homogenen Bestandteile der Dimension h und k. Das Polynom $f(x_1x_2)$ genüge den Ungleichungen

$$f(x_1x_2)>0 \quad \text{für} \quad x_1\geqq 1,\ x_2\geqq 1,$$

$$f^{(m)}(x_1x_2)>0 \quad \text{für} \quad x_1\geqq 0,\ x_2\geqq 0,\ x_1+x_2>0.$$

ω bedeute eine positive Irrationalzahl, a eine natürliche Zahl, und es seien vier Dirichletsche Reihen definiert durch

$$Z_o(s) = Z_o(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} g(x_1 x_2) f(x_1 x_2)^{-s},$$

$$Z\,(s) = Z\,(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} g(x_1 x_2) f(x_1 x_2)^{-s},$$

$$Z_a(s) = Z_a(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{a x_2} g(x_1 x_2) f(x_1 x_2)^{-s},$$

$$Z^a(s) = Z^a(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \sum_{x_1=1}^{\infty} \sum_{x_2=a x_1+1}^{\infty} g(x_1 x_2) f(x_1 x_2)^{-s}.$$

Alle vier konvergieren in der Halbebene

$$\sigma > \frac{n+2}{m}.$$

Die letzten drei Reihen befriedigen die Identität

$$\begin{aligned} &Z^a(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \\ &\quad = Z(s \mid f(x_1 x_2),\, g(x_1 x_2)) - Z_a(s \mid f(x_2 x_1),\, g(x_2 x_1)), \end{aligned}$$

während die erste den Gleichungen

$$\begin{aligned} &Z_{\frac{1}{o}}(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \\ &\quad = Z\,(s \mid f(x_1 x_2),\, g(x_1 x_2)) - Z_o(s \mid f(x_2 x_1),\, g(x_2 x_1)), \\ &Z_{a+o}(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \\ &\quad = Z_a(s \mid f(x_1 x_2),\, g(x_1 x_2)) + Z_o(s \mid f(x_1 + a x_2, x_2),\, g(x_1 + a x_2, x_2)), \\ &Z_{\frac{1}{a+o}}(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \\ &\quad = Z^a(s \mid f(x_1 x_2),\, g(x_1 x_2)) - Z_o(s \mid f(a x_1 + x_2, x_1),\, g(a x_1 + x_2, x_1)) \end{aligned}$$

genügt.

7. Die Zahl o besitze den Kettenbruch

$$o = \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots, \qquad o = \frac{1}{a_1 + o_1}, \quad o_1 = \frac{1}{a_2 + o_2}, \quad o_2 = \frac{1}{a_3 + o_3}, \ \ldots$$

sei also ein echter Bruch. Die zugehörigen Kettenbruch-Nenner und -Zähler werden durch die Rekursionsformeln

$$\begin{array}{lll} p_{-1} = 0, & p_0 = 1, & p_{\mu+1} = a_{\mu+1} p_\mu + p_{\mu-1}, \\ q_{-1} = 1, & q_0 = 0, & q_{\mu+1} = a_{\mu+1} q_\mu + q_{\mu-1} \end{array}$$

bestimmt. Mit ihrer Hilfe werde gesetzt

$$\begin{array}{ll} x_1^{(\mu)} = p_\mu x_1 + q_\mu x_2, & x_2^{(\mu)} = p_{\mu-1} x_1 + q_{\mu-1} x_2, \\ f_\mu(x_1 x_2) = f(x_1^{(\mu)} x_2^{(\mu)}), & g_\mu(x_1 x_2) = g(x_1^{(\mu)} x_2^{(\mu)}). \end{array}$$

Offenbar genügt das Polynom $f_\mu(x_1 x_2)$ als Funktion von x_1 und x_2 denselben Ungleichungen wie $f(x_1 x_2)$.

Die wiederholte Anwendung der Funktionalgleichungen in 6. führt mit den vorigen Abkürzungen zu der Identität:

$$Z_0(s|f(x_1 x_2), g(x_1 x_2)) = Z_\nu(s|f(x_1 x_2), g(x_1 x_2)) + (-1)^\nu Z_{0\nu}(s|f_\nu(x_1 x_2), g_\nu(x_1 x_2)),$$

$$Z_\nu(s) = Z_\nu(s|f(x_1 x_2),\ g(x_1 x_2)) = \sum_{\mu=0}^{\nu-1} (-1)^\mu Z^{a_{\mu+1}}(s|f_\mu(x_1 x_2),\ g_\mu(x_2 x_2)).$$

8. Die Funktion

$$Z(s|f_\mu(x_1 x_2), g_\mu(x_1 x_2)) = \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} g_\mu(x_1 x_2), f_\mu(x_1 x_2)^{-s}$$

ist von der in 1. betrachteten Art, ebenso

$$Z^a(s|f_\mu(x_1 x_2), g_\mu(x_1 x_2)) = \sum_{x_1=1}^{\infty} \sum_{x_2=a x_1+1}^{\infty} g_\mu(x_1 x_2)\, f_\mu(x_1 x_2)^{-s}$$

$$= \sum_{x_1=1}^{\infty} \sum_{x_2=1}^{\infty} g_\mu(x_1, a x_1 + x_2)\, f_\mu(x_1, a x_1 + x_2)^{-s}.$$

Die beiden Funktionen $Z_a(s)$ und $Z_\nu(s)$ lassen sich aus endlichvielen solchen Summen zusammensetzen. Man gelangt folglich zu dem Ergebnis:

Die beiden Funktionen

$$Z(s|f_\mu(x_1 x_2),\ g_\mu(x_1 x_2)), \qquad Z^a(s|f_\mu(x_1 x_2),\ g_\mu(x_1 x_2))$$

und die beiden weiteren Funktionen

$$Z_a(s|f_\mu(x_1 x_2),\ g_\mu(x_1 x_2)), \qquad Z_\nu(s|f(x_1 x_2),\ g(x_1 x_2))$$

lassen sich in die ganze Ebene meromorph fortsetzen; sie haben einfache Pole höchstens an den Stellen

$$s = \frac{n+2-k}{m}, \qquad \begin{pmatrix} k = 0,\ 1,\ 2, \ldots \\ s \neq 0,\ -1,\ -2, \ldots \end{pmatrix},$$

und sie sind für $|t| \to \infty$ gleichmäßig in σ von der Größenordnung

$$O(|t|^{2k})$$

in jedem endlichen Streifen

$$\frac{n+2-k}{m} < \sigma_1 \leqq \sigma \leqq \sigma_2 \qquad (k = 0, 1, 2, \ldots).$$

III.

9. Der Kettenbruch o sei rein periodisch, ν die Länge seiner Periode, so daß

$$o = o_\nu$$

ist. Ohne Einschränkung kann ν gerade genommen werden. Die Zahl o ist eine quadratische Irrationalität und genügt der Gleichung

$$x = \frac{q_\nu + q_{\nu-1} x}{p_\nu + p_{\nu-1} x},$$

während die andere Wurzel gleich

$$o' = \frac{q_{\nu-1} - p_\nu}{p_{\nu-1}} - o = -\frac{q_\nu}{p_{\nu-1}\,o}$$

ist. Offenbar ist o' im Gegensatz zu o negativ. Die Zahlen o und o' sind die einzigen Wurzeln des Gleichungssystems

$$x = \frac{q_\nu + q_{\nu-1}x^{(1)}}{p_\nu + p_{\nu-1}x^{(1)}},\quad x^{(1)} = \frac{q_\nu + q_{\nu-1}x^{(2)}}{p_\nu + p_{\nu-1}x^{(2)}},\ \ldots,\ x^{(k-1)} = \frac{q_\nu + q_{\nu-1}x}{p_\nu + p_{\nu-1}x},\quad (k = 1, 2, \ldots),$$

welchen Wert auch k hat.

Vermöge der besonderen Wahl von o nimmt die Funktionalgleichung für $Z_0(s)$ die Form an:

$$Z_0(s\,|\,f(x_1x_2),\, g(x_1x_2)) = Z_\nu(s\,|\,f(x_1x_2),\, g_\nu(x_1x_2)) + Z_0(s\,|\,f_\nu(x_1x_2),\, g_\nu(x_1x_2)).$$

10. Es sei speziell

$$f(x_1x_2) = \prod_{h=1}^{m}(x_1 + \varphi_h x_2),\qquad g(x_1x_2) = \prod_{k=1}^{n}(x_1 + \psi_k x_2)$$

und ferner gebe es zwei Konstante φ und ψ, so daß

$$f_\nu(x_1x_2) \equiv \varphi f(x_1x_2),\qquad g_\nu(x_1x_2) \equiv \psi g(x_1x_2)$$

ist. Dazu ist notwendig und hinreichend, daß jede der Zahlen

$$\varphi_1, \varphi_2, \ldots, \varphi_m, \psi_1, \psi_2, \ldots, \psi_m$$

einem Gleichungssystem

$$x = \frac{q_\nu + q_{\nu-1}x^{(1)}}{p_\nu + p_{\nu-1}x^{(1)}},\quad x^{(1)} = \frac{q_\nu + q_{\nu-1}x^{(2)}}{p_\nu + p_{\nu-1}x^{(2)}},\ \ldots,\ x^{(k-1)} = \frac{q_\nu + q_{\nu-1}x}{p_\nu + p_{\nu-1}x}$$

genügt. Also sind sie entweder gleich o oder gleich o'.

Wegen der Ungleichung

$$f^{(m)}(x_1x_2) > 0 \quad \text{für} \quad x_1 \geqq 0,\ x_2 \geqq 0,\ x_1 + x_2 > 0$$

kommt für $f(x_1x_2)$ nur die Wahl

$$f(x_1x_2) = (x_1 + o\,x_2)^m$$

in Frage. Dagegen darf $g(x_1x_2)$ gleich irgendeiner der $n+1$ Funktionen

$$g(x_1x_2) = (x_1 + o\,x_2)^{n-l}(x_1 + o'x_2)^l \qquad (l = 0, 1, 2, \ldots, n)$$

sein. Wegen der Identität

$$Z_0(s\,|\,(x_1 + o\,x_2)^m,\, g(x_1x_2)) = Z_0(ms\,|\,x_1 + ox_2,\, g(x_1x_2))$$

darf folglich ohne Einschränkung

$$m = 1,\quad f(x_1x_2) = x_1 + o\,x_2,\quad g(x_1x_2) = (x_1 + o\,x_2)^{n-l}(x_1 + o'x_2)^l$$

genommen werden. Die Multiplikatoren φ und ψ sind gleich

$$\varphi = \eta,\quad \psi = \eta^{n-2l},\qquad \eta = p_\nu + p_{\nu-1}o,$$

wobei die Gleichung

$$(p_\nu + p_{\nu-1}o)(p_\nu + p_{\nu-1}o') = 1$$

benutzt ist. Die Zahl η ist eine Einheit des Körpers $K(o)$.

Die Funktionalgleichung für $Z_0(s)$ nimmt die Form an:

$$Z_0(s\,|\,f(x_1 x_2), g(x_1 x_2)) = \frac{Z_\nu(s\,|\,f(x_1 x_2), g(x_1 x_2))}{1-\eta^{-s+n-2l}},$$

so daß man zu folgenden Aussagen gelangt:

Die Funktion

$$Z_0(s,l) = Z_0(s+n\,|\,x_1 + o x_2, (x_1 + o x_2)^{n-l}(x_1 + o' x_2)^l)$$

$$= \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} \left(\frac{x_1 + o' x_2}{x_1 + o x_2}\right)^l (x_1 + o x_2)^{-s}$$

gestattet die explizite Darstellung

$$Z_0(s|l) = \frac{1}{1-\eta^{-s-2l}} \cdot \sum_{\mu=0}^{\nu-1} (-1)^\mu \frac{(p_\mu + o' p_{\mu+1})^l}{(p_\mu + o p_{\mu+1})^{s+l}} \cdot Z\left(s, l \,\middle|\, \frac{q_\mu + o q_{\mu+1}}{p_\mu + o p_{\mu+1}}, \frac{q_\mu + o' q_{\mu+1}}{p_\mu + o' p_{\mu+1}}\right).$$

Sie läßt sich also in die ganze Ebene als meromorphe Funktion fortsetzen; ihre Pole sind einfach und liegen in den Punkten

$$s = 2, 1, 0, \ldots, 2-l,$$

$$s = -2l + \frac{2k^* \pi i}{\log \eta} \qquad (k^* = 0, \mp 1, \mp 2, \ldots).$$

Bleibt s von diesen Punkten um mehr als eine feste positive Strecke entfernt, so gilt gleichmäßig in s und l im Streifen

$$2 - k < \sigma_1 \leqq \sigma \leqq \sigma_2 \qquad (k = 0, 1, 2, \ldots)$$

die Ungleichung:

$$|Z_0(s,l)| \leqq \sum_{\mu=0}^{\nu-1} c_\mu \frac{(p_\mu + |o'|\,p_{\mu+1})^l}{(p_\mu + o p_{\mu+1})^{l+\sigma}} \left\{ 2\, \frac{1 + \left|\dfrac{q_\mu + o' q_{\mu+1}}{p_\mu + o' p_{\mu+1}}\right|}{\min\left(1, \dfrac{q_\mu + o q_{\mu+1}}{p_\mu + o p_{\mu+1}}\right)} \right\}^l (1+|t|)^{2k}$$

mit ν Konstanten

$$c_0, c_1, \ldots, c_{\nu-1} > 0,$$

die von s und l nicht abhängen. Es gibt also auch zwei Konstante

$$C > 0, \quad \Gamma > 0,$$

die von s und l nicht abhängen, so daß

$$|Z_0(s,l)| \leqq \Gamma C^l (1+|t|)^{2k}$$

ist, unter den vorigen Bedingungen.

IV.

11. Die Funktion $f(x_1 x_2)$ sei jetzt gleich

$$f(x_1 x_2) = \prod_{h=1}^{m} (o_h^{(1)} x_1 + o_h^{(2)} x_2),$$

also in eine Form ausgeartet. Die Quotienten $\frac{o_h^{(1)}}{o_h^{(2)}}$ sind entweder positiv oder zu je zweien konjugiert komplex. Es ist

$$f_N(x_1 x_2) = \prod_{h=1}^{m} \left\{ (o_h^{(1)} p_N + o_h^{(2)} p_{N-1}) x_1 + (o_h^{(1)} q_N + o_h^{(2)} q_{N-1}) x_2 \right\}.$$

Wenn erstens $\frac{o_h^{(1)}}{o_h^{(2)}}$ positiv ist, so gilt

$$\lim_{N \to \infty} \frac{o_h^{(1)} q_N + o_h^{(2)} q_{N-1}}{o_h^{(1)} p_N + o_h^{(2)} p_{N-1}} = o$$

wegen

$$\lim_{N \to \infty} \frac{q_N}{p_N} = \lim_{N \to \infty} \frac{q_{N-1}}{p_{N-1}} = c.$$

Zweitens sei $\frac{o_h^{(1)}}{o_h^{(2)}}$ nichtreell:

$$o_h^{(1)} \neq 0, \quad o_h^{(2)} \neq 0, \quad o_h^{(2)} = (\alpha + \beta_i) o_h^{(1)}, \quad \beta \neq 0;$$

dann ist

$$\Omega_N = \frac{o_h^{(1)} q_N + o_h^{(2)} q_{N-1}}{o_h^{(1)} p_N + o_h^{(2)} p_{N-1}} = \frac{q_N + (\alpha + \beta i) q_{N-1}}{p_N + (\alpha + \beta i) p_{N-1}}$$

$$= \frac{\frac{q_N}{p_N} \cdot \frac{p_N}{p_{N-1}} + (\alpha + \beta i) \frac{q_{N-1}}{p_{N-1}}}{\frac{p_N}{p_{N-1}} + (\alpha + \beta i)} = \frac{q_{N-1}}{p_{N-1}} \cdot \frac{\left(\frac{q_N}{p_N} \cdot \frac{p_{N-1}}{q_{N-1}} \right) \frac{p_N}{p_{N-1}} + (\alpha + \beta i)}{\frac{p_N}{p_{N-1}} + (\alpha + \beta i)},$$

ferner, wenn N ins Unendliche strebt:

$$\frac{q_{N-1}}{p_{N-1}} = o + \mathrm{o}(1), \quad \frac{p_{N-1}}{q_{N-1}} = \frac{1}{o} + \mathrm{o}(1), \quad \frac{q_N}{p_N} = o + \mathrm{o}(1), \quad \frac{p_N}{p_{N-1}} = \mathrm{O}(1)$$

und

$$\left| \frac{1}{\frac{p_N}{p_{N-1}} + (\alpha + \beta i)} \right| \leq \frac{1}{|\beta|} = \mathrm{O}(1).$$

Also hat man

$$\Omega_N = (o + o(1)) \frac{(o + o(1))\left(\frac{1}{o} + o(1)\right)\frac{p_N}{p_{N-1}} + (\alpha + \beta i)}{\frac{p_N}{p_{N-1}} + (\alpha + \beta i)}$$

$$= (o + o(1)) \left\{ 1 + \frac{\frac{p_N}{p_{N-1}} o(1)}{\frac{p_N}{p_{N-1}} + (\alpha + \beta i)} \right\} = (o + o(1))(1 + o(1))$$

und auch in diesem Fall

$$\lim_{N \to \infty} \frac{o_h^{(1)} q_N + o_h^{(2)} q_{N-1}}{o_h^{(1)} p_N + o_h^{(2)} p_{N-1}} = o.$$

In der Darstellung

$$f_N(x_1, x_2) = A_N F_N(x_1 x_2); \qquad A_N = \prod_{h=1}^{m} (o_h^{(1)} p_N + o_h^{(2)} p_{N-1})$$

ist daher

$$F_N(x_1 x_2) = (x_1 + o x_2)^m + \mathfrak{F}_N(x_1 x_2)$$

und sämtliche Koeffizienten der Form $\mathfrak{F}_N(x_1 x_2)$ werden für genügend großes N beliebig klein.

12. Wir nehmen jetzt für

$$f(x_1 x_2) = \prod_{h=1}^{m} (o_h^{(1)} x_1 + o_h^{(2)} x_2)$$

die vorige Funktion an, für $g(x_1 x_2)$ die Funktion

$$g(x_1 x_2) = (x_1 + o x_2)^{n-l} (x_1 + o' x_2)^l$$

und ferner N als ein sehr großes Vielfaches von ν:

$$\mathrm{N} = N \nu.$$

Dann ist

$$Z_0(s \mid f(x_1 x_2),\, g(x_1 x_2)) = Z_N(s \mid f(x_1 x_2),\, g(x_1 x_2)) +$$
$$+ A_N^{-s} \eta^{N(n-2l)} Z_0(s \mid (x_1 + o x_2)^m + \mathfrak{F}_N(x_1 x_2),\, g(x_1 x_2))$$

und folglich

$$Z_0(s \mid f(x_1 x_2),\, g(x_1 x_2)) = Z_N(s \mid f(x_1 x_2),\, g(x_1 x_2))$$
$$+ A_N^{-s} \eta^{N(n-2l)} \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} g(x_1 x_2)(x_1 + o x_2)^{-ms} \left\{1 + \frac{\mathfrak{F}_N(x_1 x_2)}{(x_1 + o x_2)^m}\right\}^{-s}$$

Bedeutet ϑ eine beliebig kleine positive Konstante und ist N oberhalb einer nur von ϑ abhängigen Zahl, so gilt gleichmäßig für alle natürlichen Werte von x_1 und x_2:

$$\left| \frac{\mathfrak{F}_N(x_1 x_2)}{(x_1 + o x_2)^m} \right| \leq \vartheta$$

und folglich auch die binomische Reihe

$$\left\{1+\frac{\mathfrak{F}_N(x_1 x_2)}{(x_1+o x_2)^m}\right\}^{-s}=\sum_{\lambda=0}^{\infty}\binom{-s}{\lambda}\frac{\mathfrak{F}_N(x_1 x_2)^\lambda}{(x_1+o x_2)^{\lambda m}}.$$

Eine erlaubte Summationsvertauschung führt daher zu der Entwicklung:

$$Z_o(s|(x_1+o x_2)^m+\mathfrak{F}_N(x_1 x_2),\ g(x_1 x_2))=$$

$$=\sum_{\lambda=0}^{\infty}\binom{-s}{\lambda}Z_o(m(s+\lambda)\,|\,x_1+o x_2,\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^\lambda),$$

die zunächst für

$$\sigma>\frac{n+2}{m}$$

Gültigkeit hat.

13. Die $m+1$ Formen

$$(x_1+o x_2)^{m-\mu}(x_1+o' x_2)^\mu \qquad (\mu=0,1,2,\ldots,m)$$

sind offenbar linear unabhängig in bezug auf den Körper der komplexen Zahlen. Die Form $\mathfrak{F}_N(x_1 x_2)$ läßt sich aus ihnen in der Gestalt

$$\mathfrak{F}_N(x_1 x_2)=\sum_{\mu=0}^{m}a_\mu\binom{m}{\mu}(x_1+o x_2)^{m-\mu}(x_1+o' x_2)^\mu,\quad |a_\mu|\leqq\theta$$

darstellen, wo die positive Konstante θ zugleich mit ϑ gegen Null strebt. Folglich wird

$$\mathfrak{F}_N(x_1 x_2)^\lambda=\sum_{\mu=0}^{m\lambda}a_\mu^{(\lambda)}\binom{m\lambda}{\mu}(x_1+o x_2)^{m\lambda-\mu}(x_1+o' x_2)^\mu, \qquad |a_\mu^{(\lambda)}|\leqq\theta^\lambda,$$

$$g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^\lambda=\sum_{\mu=0}^{m\lambda}a_\mu^{(\lambda)}\binom{m\lambda}{\mu}(x_1+o x_2)^{m\lambda+n-\mu-l}(x_1+o' x_2)^{\mu+l}$$

Man hat damit

$$Z_o(m(s+\lambda)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^\lambda)=\sum_{\mu=0}^{m\lambda}a_\mu^{(\lambda)}\binom{m\lambda}{\mu}Z_o(ms-n,\ \mu+l)$$

und aus 10. ergibt sich:

Die Funktion

$$Z_o(m(s+\lambda)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^\lambda)$$

läßt sich als meromorphe Funktion in die ganze Ebene fortsetzen. Sie hat einfache Pole in den Punkten

$$s=\frac{n+2-k}{m} \qquad (k=0,1,\ldots,\lambda m+l),$$

$$s=\frac{n-2k}{m}+\frac{2k^*\pi i}{m\log\eta} \quad \begin{pmatrix}k=l,\,l+1,\ldots,\,l+\lambda m\\ k^*=0,\mp 1,\mp 2,\ldots\end{pmatrix}.$$

Bleibt s von diesen Punkten um mehr als eine feste positive Strecke entfernt, so gilt gleichmäßig in s und λ im Streifen

$$\frac{n+2-k}{m}<\sigma_1\leqq\sigma\leqq\sigma_2$$

die Ungleichung

$$|Z_0(m(s+\lambda)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^{\lambda})| \leqq$$

$$\leqq \sum_{\mu=0}^{\lambda m}\theta^{\lambda}\binom{m\lambda}{\mu}\Gamma C^{\mu+l}(1+|t|)^{k} = \Gamma C^{l}(1+|t|)^{k}\{\theta(1+C)^{m}\}^{\lambda}.$$

14. Das vorige Ergebnis läßt sich unmittelbar auf die Funktion

$$Z_0(s\,|\,f(x_1 x_2),\ g(x_1 x_2)) = Z_N(s\,|\,f(x_1 x_2),\ g(x_1 x_2)) +$$

$$+ A_N^{-s}\eta^{N(n-2l)}\sum_{\lambda=0}^{\infty}\binom{-s}{\lambda}Z_0(m(s+l)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^{\lambda})$$

anwenden; man erhält den Satz:

Die Funktion

$$Z_0(s\,|\,f(x_1 x_2),\ g(x_1 x_2)),\quad \text{mit}\quad \begin{aligned} f(x_1 x_2) &= \prod_{k=1}^{m}(o_k^{(1)}x_1 + o_k^{(2)}x_2),\\ g(x_1 x_2) &= (x_1+o x_2)^{n-l}(x_1+o' x_2)^{l}, \end{aligned}$$

läßt sich als meromorphe Funktion in die ganze Ebene fortsetzen. Sie hat einfache Pole allein in den Punkten

$$s = \frac{n+2-k}{m} \qquad (k=0,1,\ldots)$$

$$s = \frac{n-2k}{m} + \frac{2k^{*}\pi i}{m\log\eta} \qquad \begin{pmatrix} k = 0,1,\ldots \\ k^{*} = 0, \mp 1, \mp 2, \ldots \end{pmatrix}$$

Es sei ε eine beliebig kleine positive Konstante; s bleibe von den vorigen Polen um mehr als eine feste positive Strecke entfernt. Dann gilt gleichmäßig in s in jedem endlichen Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2$$

die Ungleichung

$$Z_0(s\,|\,f(x_1 x_2),\ g(x_1 x_2)) = O(e^{\varepsilon|t|}).$$

Man hat nämlich für $\frac{n+2-k}{m} < \sigma_1$ und $k = 0, 1, \ldots$:

$$\left|\sum_{\lambda=0}^{\infty}\binom{-s}{\lambda}Z_0(m(s+l)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^{\lambda})\right| \leqq$$

$$\leqq \sum_{\lambda=0}^{\infty}(-1)^{\lambda}\binom{-|s|}{\lambda}\Gamma C^{l}(1-|t|)^{k}\{\theta(1+C)^{m}\}^{\lambda} \leqq$$

$$\leqq \Gamma C^{l}(1+|t|)^{k}\{1-\theta(1+C)^{m}\}^{-|s|}.$$

Wenn man N hinreichend groß nimmt, so wird aber ϑ und damit θ beliebig klein; man kann durch geeignete Wahl von N daher erreichen, daß

$$\theta(1+C)^{m} \leqq 1 - e^{-\frac{\varepsilon}{2}}.$$

Dann ist

$$\left|\sum_{\lambda=0}^{\infty}\binom{-s}{\lambda} Z_0(m(s+l)\,|\,(x_1+o x_2),\ g(x_1 x_2)\,\mathfrak{F}_N(x_1 x_2)^{\lambda})\right| \leqq$$

$$\leqq \Gamma C^l(1+|t|^k)e^{\frac{\varepsilon}{2}|t|} = O(e^{\varepsilon|t|}).$$

Da ferner nach 8. die Funktion $\mathsf{Z}_N(s)$ im Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2$$

nur von der Größenordnung einer Potenz von $1+|t|$ ist, so folgt die Behauptung.

V.

15. Nunmehr sei o eine beliebige quadratische Irrationalzahl; das Polynom $f(x_1 x_2)$ genüge allein den früheren Ungleichungen und $g(x_1 x_2)$ sei ganz beliebig. Auch in diesem Fall kann die Fortsetzbarkeit von $Z_0(s)$ bewiesen werden.

Es bedeutet jedoch keine Einschränkung, wenn der Kettenbruch von o reinperiodisch angenommen wird. Denn als Kettenbruch einer quadratischen Irrationalität ist er gewiß periodisch; die Funktionalgleichungen in 6. und 7. gestatten, den unperiodischen Teil des Kettenbruches abzuschneiden. Aus dem gleichen Grund darf man o als echten Bruch annehmen.

Weiter darf ohne Einschränkung $g(x_1 x_2)$ als eine Form

$$(x_1+o x_2)^{n-l}(x_1+o' x_2)^l$$

angenommen werden. Denn andernfalls kann $g(x_1 x_2)$ in der Form

$$g(x_1 x_2) = \sum_{\alpha=0}^{n}\sum_{\beta=0}^{\alpha} A_{\alpha\beta}(x_1+o x_2)^{\alpha-\beta}(x_1+o' x_2)^{\beta}$$

mit konstanten Koeffizienten $A_{\alpha\beta}$ dargestellt werden, so daß

$$Z_0(s\,|\,f(x_1 x_2),\ g(x_1 x_2)) = \sum_{\alpha=0}^{n}\sum_{\beta=0}^{\alpha} A_{\alpha\beta} Z_0(s\,|\,f(x_1 x_2),\ (x_1+o x_2)^{\alpha-\beta}(x_1+o' x_2)^{\beta})$$

wird, und man also $Z_0(s\,|\,f(x_1 x_2),\ g(x_1 x_2))$ als Summe endlichvieler $Z_0(s)$-Funktionen erhält, wo jetzt $g(x_1 x_2)$ eine Form der angegebenen Art ist.

16. Sei also

a: der Kettenbruch von o reinperiodisch: $o = o_\nu$ und ν gerade;

b: das Polynom $g(x_1 x_2)$ gleich der Form $g(x_1 x_2) = (x_1+o x_2)^{n-l}(x+o' x_2)^l$;

c: das Polynom $f(x_1 x_2)$ den früheren Ungleichungen genügend und sonst beliebig.

Die Menge aller Paare natürlicher Zahlen (x_1, x_2) mit

$$1 \leqq x_1 \leqq o x_2$$

sei $\mathfrak{T}$, die Teilmenge der Paare mit

$$\left|\frac{r(x_1 x_2)}{f^{(m)}(x_1 x_2)}\right| < \varepsilon, \qquad r(x_1 x_2) = f(x_1 x_2) - f^{(m)}(x_1 x_2)$$

heiße $\mathfrak{T}_1$, die Komplementärmenge $\mathfrak{T} - \mathfrak{T}_1$ werde mit $\mathfrak{T}_2$ bezeichnet. Die Menge $\mathfrak{T}_2$ enthält offenbar nur endlichviele Elemente. Dabei ist ε eine positive beliebig kleine Zahl.

In $\mathfrak{T}_1$ besteht die Binomialreihe

$$\left(1 + \frac{r(x_1 x_2)}{f^{(m)}(x_1 x_2)}\right)^{-s} = \sum_{\mu=0}^{\infty} \binom{-s}{\mu} \left(\frac{r(x_1 x_2)}{f^{(m)}(x_1 x_2)}\right)^{\mu};$$

folglich ist für $\sigma > \frac{n+2}{m}$:

$$\sum_{\mathfrak{T}} g(x_1 x_2) f(x_1 x_2)^{-s} =$$

$$= \sum_{\mathfrak{T}_2} g(x_1 x_2) f(x_1 x_2)^{-s} + \sum_{\mu=0}^{\infty} \binom{-s}{\mu} \sum_{\mathfrak{T}_1} g(x_1 x_2) r(x_1 x_2)^{\mu} f^{(m)}(x_1 x_2)^{-(s+\mu)}.$$

Wird aber s auf einen beliebigen endlichen Bereich beschränkt, so ist gleichmäßig in s für große μ:

$$\sum_{\mathfrak{T}_1} g(x_1 x_2) r(x_1 x_2)^{\mu} f^{(m)}(x_1 x_2)^{-(s+\mu)} = O(\varepsilon^{\mu}).$$

Die Reihenentwicklung

$$Z_o(s \mid f(x_1 x_2),\, g(x_1 x_2)) = \sum_{\mathfrak{T}_2} g(x_1 x_2) f(x_1 x_2)^{-s} +$$

$$+ \sum_{\mu=0}^{\infty} \binom{-s}{\mu} \left\{ Z_o(s+\mu \mid f^{(m)}(x_1 x_2),\, g(x_1 x_2) r(x_1 x_2)^{\mu}) - \sum_{\mathfrak{T}_2} g(x_1 x_2) r(x_1 x_2)^{\mu} f^{(m)}(x_1 x_2)^{-(s+\mu)} \right\}$$

konvergiert somit in jedem endlichen Bereich, der keine Pole der Summanden enthält, gleichmäßig in s. Weiter ist in jedem endlichen Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2,$$

wenn s von den Polen der Summanden um mehr als eine feste positive, beliebig kleine Strecke entfernt bleibt, gleichmäßig in s für große μ:

$$Z_o(s+\mu \mid f^{(m)}(x_1 x_2),\, g(x_1 x_2) r(x_1 x_2)^{\mu}) - \sum_{\mathfrak{T}} g(x_1 x_2) r(x_1 x_2)^{\mu} f^{(m)}(x_1 x_2)^{-(s+\mu)} = O(\varepsilon^{\mu});$$

$$\sum_{\mu=0}^{\infty} \left| \binom{-s}{\mu} \right| \varepsilon^{\mu} = O(e^{\varepsilon |t|}).$$

Es folgt also nach 14.:

Seien $f(x_1 x_2)$ und $g(x_1 x_2)$ zwei Polynome vom Grad m und n in den x mit reellen Koeffizienten; $f(x_1 x_2)$ genüge den Ungleichungen

$$f(x_1 x_2) > 0 \text{ für } x_1 \geqq 1,\ x_2 \geqq 1;\quad f^{(m)}(x_1 x_2) > 0 \text{ für } x_1 \geqq 0,\ x_2 \geqq 0,\ x_1 + x_2 > 0.$$

Die natürliche Zahl d gleich 2, wenn $r(x_1 x_2)$ identisch verschwindet, sonst gleich 1.

o sei eine positive quadratische Irrationalzahl, ζ die erzeugende positive Einheit des Zahlkörpers $K(o)$.

Die durch die Reihe

$$Z_o(s \mid f(x_1 x_2),\ g(x_1 x_2)) = \sum_{x_2=1}^{\infty} \sum_{x_1=1}^{[o x_2]} g(x_1 x_2)\, f(x_1 x_2)^{-s}$$

definierte Funktion läßt sich meromorph in die ganze Ebene fortsetzen; sie hat einfache Pole allein an den Stellen

$$\begin{aligned} s &= \frac{n+2-k}{m} \\ s &= \frac{n-dk}{m} + \frac{2k^* \pi i}{m \log \eta} \end{aligned} \qquad \begin{pmatrix} k = 0, 1, 2, \ldots \\ k^* = 0, \mp 1, \mp 2, \ldots \end{pmatrix}.$$

Dabei bedeutet η eine geeignete ganze Potenz von ζ.

Bleibt s von den Polen um mehr als eine feste positive, beliebig kleine Strecke entfernt, so gilt im Streifen

$$\sigma_1 \leqq \sigma \leqq \sigma_2$$

gleichmäßig in s:

$$Z_o(s \mid f(x_1 x_2),\ g(x_1 x_2)) = O(e^{\varepsilon |t|}).$$

17. Dem vorigen Satz haftet noch das Unbefriedigende an, daß Annahmen über $f(x_1 x_2)$ gemacht werden in einem Bereich, der über das Summationsgebiet hinausgeht. Es läßt sich aber zeigen, daß der vorige Satz auch noch bestehen bleibt, wenn man z. B. folgende Annahmen macht:

„Es sei ω eine beliebige positive Zahl, die größer als o ist; das Polynom $f(x_1 x_2)$ genüge den Ungleichungen:

$$f(x_1 x_2) > 0 \quad \text{für} \quad 0 \leqq x_1 \leqq \omega x_2,$$
$$f^{(m)}(x_1 x_2) > 0 \quad \text{für} \quad 0 \leqq x_1 \leqq \omega x_2,\quad x_1 + x_2 > 0.$$

Dann bleibt der vorige Satz bestehen."

Krefeld, 14. 9. 1928.

(Eingegangen am 28. 10. 1928.)

Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren.

Von

J. v. Neumann in Berlin.

Einleitung.

I. Wenn wir alle Funktionen eines gewissen metrischen Raumes Ω (z. B. der reellen Zahlengeraden, der komplexen Zahlenebene, der Oberfläche der Einheitskugel, der Strecke $0, 1$ usw.) betrachten, die gewissen Regularitätsbedingungen genügen (z. B. stetig und bis auf endlich viele Knicke stetig differentiierbar sind, zweimal stetig differentiierbar sind, ein endliches Absolutwertquadratintegral über Ω haben[1]) usw.) und eventuell auch noch gewissen Randbedingungen unterworfen sind, so wird unter einem linearen Operator bekanntlich das Folgende verstanden: Eine Zuordnung R, die jeder Funktion f unserer Klasse eine Funktion Rf (die nicht mehr zu dieser Klasse gehören muß) zuordnet, aber so, daß stets

$R(a_1 f_1 + \dots + a_k f_k) = a_1 R f_1 + \dots + a_k R f_k$ ($a_1, \dots, a_k$ komplexe Konstante) ist.

Wenn in Ω ein allgemeiner Maßbegriff (etwa im Sinne des Lebesgueschen) existiert und dv das Volumelement in Ω ist (auf der Geraden: dx, in der Ebene: $dx\,dy$, auf der Oberfläche der Einheitskugel: $\sin\vartheta \cdot d\vartheta\, d\varphi$ usw.), und das Integral über diesen ganzen Raum mit $\int\limits_{\Omega}$ bezeichnet wird, so heißt ein linearer Operator R selbstadjungiert oder Hermitesch, wenn für alle zugelassenen Funktionen f, g

$$\int\limits_{\Omega} f \cdot \overline{Rg} \cdot dv = \int\limits_{\Omega} Rf \cdot \bar{g} \cdot dv$$

gilt[2]).

---

[1]) Die Funktionen dürfen komplexe Werte haben.

[2]) $\bar{f}$ ist diejenige Funktion, die überall den zum Werte von f konjugiert-komplexen Wert annimmt. — Unsere Terminologie weicht von der üblichen etwas ab:

(Fortsetzung der Fußnote [2]) auf nächster Seite.)

Hermitesche lineare Operatoren (wir wollen sie kurz H. O. nennen) sind z. B. die folgenden (wir geben nun auch die Argumente der Funktionen an, und zwar sei P der allgemeine Punkt von Ω):

$R f(P) = f(P)$

$R f(P) = x \cdot f(P)$ (x irgendeine Koordinate von P in Ω),

$R f(P) = \int\limits_{\Omega} \varphi(P', P) \cdot f(P') \cdot dv'$ ($\varphi(P', P) = \overline{\varphi(P, P')}$, d. h. der Kern Hermitesch),

$R f(P) = L f(P)$ (L irgendein selbstadjungierter Differential- oder partieller Differentialausdruck, vgl. Anm. [2])).

Bekanntlich bestimmen die beiden letzten Typen von H. O. ein Eigenwertproblem, das bei hinreichender Regularität des Kernes $\varphi(P, P')$ bzw. der Koeffizienten von L (nämlich wenn nur ein „Punktspektrum" existiert) so formuliert werden kann[3]):

Ein Eigenwert ist eine Zahl λ, zu der es eine Funktion $f \neq 0$ mit

$$R f = \lambda f$$

gibt; f ist dann Eigenfunktion. Es gibt im allgemeinen unendlich viele Eigenwerte $\lambda_1, \lambda_2, \ldots$, denen Eigenfunktionen $\varphi_1, \varphi_2, \ldots$ so zugeordnet werden können, daß sie ein vollständiges System von Orthogonalfunktionen bilden. D. h. es gilt:

$$\int\limits_{\Omega} \varphi_m \cdot \overline{\varphi_n} \cdot dv = \begin{cases} 0, & \text{für } m = n \\ 1, & \text{für } m \neq n \end{cases} \quad \text{(Orthogonalität)},$$

ein Differentialoperator R z. B. heißt ja gewöhnlich dann selbstadjungiert, wenn $f \cdot \overline{Rg} - Rf \cdot g$ vollständiges Differential eines Ausdruckes der f, g und ihrer Ableitungen ist, also

$$\int\limits_{\Omega} f \cdot \overline{Rg} \cdot dv - \int\limits_{\Omega} Rf \cdot \bar{g} \cdot dv$$

ein nur von den Randwerten von f, g und ihrer Ableitungen abhängiger Ausdruck. Unsere Selbstadjungiertheit (Hermitescher Charakter) folgt hieraus nur, wenn dieser Ausdruck, infolge geeigneter Randbedingungen, verschwindet. D. h. ein im gewöhnlichen Sinne selbstadjungierter Differentialoperator ist es für uns eventuell erst in einem hinreichend eingeengten Definitionsbereiche (vgl. auch Einleitung VII).

[3]) Das Eigenwertproblem der regulären Integraloperatoren wurde von Hilbert gelöst (Göttinger Nachr. 1904, S. 49—91), seine Methode wurde später von Courant, Hellinger, F. Riesz, E. Schmidt, Toeplitz u. a. bedeutend vereinfacht. Die Differentialoperatoren (zumindest die selbstadjungierten von zweitem Grade) können bei der Lösung des Eigenwertproblems durch Integraloperatoren ersetzt werden (Hilbert, Göttinger Nachr. 1904, S. 213—259), die dieselben Eigenfunktionen, aber reziproke Eigenwerte haben.

und für jedes Paar zugelassener Funktionen f, g

$$\int_\Omega f\cdot\bar{g}\cdot dv = \sum_{n=1}^{\infty}\left(\int_\Omega f\cdot\overline{\varphi_n}\cdot dv\right)\cdot\overline{\left(\int_\Omega g\cdot\overline{\varphi_n}\cdot dv\right)} \mid \text{(Vollständigkeit)}.$$

Wie für alle vollständigen Orthogonalsysteme, gilt der Entwicklungssatz mit Mittelkonvergenz:

$$\int_\Omega \left| f - \sum_{n=1}^{N} a_n\cdot\varphi_n\right|^2\cdot dv \to 0 \qquad \left(a_n = \int_\Omega f\cdot\overline{\varphi_n}\cdot dv\right)$$

für $N\to\infty$.

Wenn sich $\varphi(P, P')$ bzw. die Koeffizienten von L nicht mehr ganz regulär verhalten, so treten Verwicklungen auf, die man durch das Auftreten eines Streckenspektrums kennzeichnet. Immerhin können die Begriffe Eigenwert und Eigenfunktion durch geeignete Verallgemeinerungen (insbesondere durch Herabsetzung der Regularitätsanforderungen an die Eigenfunktion oder gar durch Betrachtung sogenannter differentieller Eigenfunktionen[4]) aufrechterhalten werden. Auch die Vollständigkeitssätze bleiben bei geeigneter Verallgemeinerung (Ersetzen der Summe durch ein Integral usw.) gewahrt; aber der ganze Aufbau wird wesentlich komplizierter und verliert viel von seiner ursprünglichen Anschaulichkeit[5]).

II. Wenn wir die in I. betrachteten kontinuierlichen Räume Ω durch den „diskreten Raum" 1, 2, 3, ... der positiven ganzen Zahlen ersetzen, wobei die Funktionen $f(P)$ (P durchläuft Ω) durch die Folgen a_n ($n = 1, 2, \ldots$) und die Integrale $\int_\Omega f\cdot dv$ durch die Summen $\sum_{n=1}^{\infty} a_n$ sinngemäß zu ersetzen sind, so gelangen wir zu ganz analogen Begriffsbildungen wie in I.

Die einfachsten Beispiele von H. O. (ja, wie man nach präziserer Fassung des Begriffes zeigen kann, die einzigen) sind die linearen Transformationen (unendlich vieler Variablen) mit Hermitescher Matrix:

$$R(x_1, x_2, \ldots) = (y_1, y_2, \ldots),$$

$$y_n = \sum_{m=1}^{\infty} a_{mn} x_m \quad (n = 1, 2, \ldots), \qquad a_{mn} = \overline{a_{nm}}.$$

(Die Rolle der „Regularitätsbedingungen" bei Funktionen in Ω spielen da

---

[4]) Hellinger, Göttinger Inaugural-Dissertation 1907, S. 7; Carleman, Sur les équations intégrales à noyau réel et symétrique, Upsala 1923, S. 82.

[5]) Das Eigenwertproblem der Integralgleichungen mit „beschränktem" Kerne wurde von Weyl gelöst (Göttinger Inaugural-Dissertation 1908) mit Hilfe der Hilbertschen Theorie der beschränkten Bilinearformen (Göttinger Nachr. 1906, S. 157—209). Auch für eine ausgedehnte Klasse selbstadjungierter Differentialoperatoren zweiten Grades, deren Koeffizienten singulär sein dürfen, erledigte Weyl das Eigenwertproblem (Math. Annalen 68 (1910), S. 220—269). Eine noch allgemeinere Klasse von Integraloperatoren untersuchte Carleman (vgl. Anm. [4])).

Bedingungen von der Art, daß alle $\sum_{m=1}^{\infty} a_{mn} x_m$ konvergieren sollen, u. ä. Wir werden diese Matrizen noch im Anhang III eingehender betrachten.) Diese linearen Transformationen sind offenbar Analoga der Integralkern-Transformationen in kontinuierlichen Räumen Ω:

$$R f(P) = \int_{\Omega} \varphi(P', P) \cdot f(P') \cdot dv'.$$

Auch für diese Transformationen bzw. für die zu ihnen gehörigen Hermiteschen Formen

$$\sum_{m,n=1}^{\infty} a_{mn} x_m \overline{y_n}$$

ist das Eigenwertproblem von Hilbert gestellt und gelöst worden, falls die Matrix $\{a_{mn}\}$ (bzw. die Hermitesche Form $\sum_{m,n=1}^{\infty} a_{mn} x_m \overline{y_n}$) gewissen Bedingungen genügt, die der „Regularität" des Kernes $\varphi(P, P')$ bei Integralkern-Transformationen entsprechen.

Es treten dieselben Erscheinungen auf: Bei hoher Regularität (Vollstetigkeit) bloß Punktspektrum, mit denselben Eigenschaften wie in I., nur daß $\int_{\Omega}$ durch $\sum_{n=1}^{\infty}$ zu ersetzen ist. Also: λ ist Eigenwert, wenn es eine Folge $(x_1, x_2, \ldots)$ mit

$$R(x_1, x_2, \ldots) = \lambda(x_1, x_2, \ldots), \quad \text{d. h.} \ \sum_{m=1}^{\infty} a_{mn} x_m = \lambda x_n \quad (n = 1, 2, \ldots)$$

gibt [ohne daß $(x_1, x_2, \ldots) = (0, 0, \ldots)$ ist, außerdem ist als „Regularitätsbedingung" die Endlichkeit von $\sum_{n=1}^{\infty} |x_n|^2$ zu fordern], und $(x_1, x_2, \ldots)$ ist dann Eigenfolge. Sind $\lambda_1, \lambda_2, \ldots$ alle Eigenwerte, so können ihnen Eigenfolgen $(x_{11}, x_{12}, \ldots), (x_{21}, x_{22}, \ldots), \ldots$ so zugeordnet werden, daß sie ein vollständiges Orthogonalsystem bilden:

$$\sum_{m=1}^{\infty} x_{pm} \overline{x_{qm}} = \begin{cases} 1, & \text{für } p = q \\ 0, & \text{für } p \neq q \end{cases} \quad \text{(Orthogonalität)},$$

und[6])

$$\sum_{m=1}^{\infty} x_{mp} \overline{x_{mq}} = \begin{cases} 1, & \text{für } p = q \\ 0, & \text{für } p \neq q \end{cases} \quad \text{(Vollständigkeit)}.$$

Bei geringerer Regularität (Beschränktheit) kann wieder ein Strecken-

---

[6]) Dies ist bekanntlich dem eigentlichen Analogon der Vollständigkeitsrelation (vgl. I.)

$$\sum_{n=1}^{\infty} a_n \overline{b_n} = \sum_{p=1}^{\infty} \left(\sum_{n=1}^{\infty} a_n \overline{x_{pn}}\right) \overline{\left(\sum_{n=1}^{\infty} b_n \overline{x_{pn}}\right)}$$

gleichbedeutend.

spektrum auftreten, was dieselben Komplikationen bedingt, wie die unter I. für kontinuierliche Räume skizzierten; da wir uns später noch recht eingehend damit zu beschäftigen haben werden, wollen wir es hier nicht weiter erörtern[7]). Irgendwelche „Regularitäts"-Annahmen wurden übrigens in der bisherigen Literatur (sowohl bei Matrizen als auch bei Integraloperatoren) stets gemacht: eine Diskussion des Eigenwertproblems auf Grund des Hermiteschen Charakters allein erfolgte nie.

III. Das analoge Verhalten der kontinuierlichen Räume Ω untereinander, sowie mit dem hier betrachteten diskreten Raume $1, 2, \ldots$, liegt bekanntlich an folgendem:

Wenn $\varphi_1, \varphi_2, \ldots$ ein vollständiges System von Orthogonalfunktionen im betrachteten Raume Ω ist, so wird durch die Zuordnung

$$f \longleftrightarrow (x_1, x_2, \ldots), \qquad x_n = \int_\Omega f \cdot \overline{\varphi_n} \cdot dv \qquad (n = 1, 2, \ldots)$$

jedem f mit endlichem Absolutwertquadratintegral $\int_\Omega |f|^2 \cdot dv$ eine Folge $(x_1, x_2, \ldots)$ mit endlicher Absolutwertquadratsumme $\sum_{n=1}^{\infty} |x_n|^2$ zugeordnet. Dabei entsprechen verschiedenen f verschiedene $(x_1, x_2, \ldots)$, und nach dem Satz von Fischer und F. Riesz[8]) entspricht wirklich jedes $(x_1, x_2, \ldots)$ mit endlichem $\sum_{n=1}^{\infty} |x_n|^2$ einem f mit endlichem $\int_\Omega |f|^2 \cdot dv$.

Die Zuordnung ist linear, d. h.

$$\text{aus} \quad f \longleftrightarrow (x_1, x_2, \ldots) \quad \text{folgt} \quad af \longleftrightarrow (a x_1, a x_2, \ldots),$$

$$\text{aus} \quad \left\{ \begin{array}{l} f \longleftrightarrow (x_1, x_2, \ldots) \\ g \longleftrightarrow (y_1, y_2, \ldots) \end{array} \right\} \quad \text{folgt} \quad f + g \longleftrightarrow (x_1 + y_1, x_2 + y_2, \ldots),$$

und es ist

$$\int_\Omega f \cdot \bar{g} \cdot dv = \sum_{n=1}^{\infty} x_n \overline{y_n} \text{ [9])}.$$

D. h. alle bei der Beschreibung der Eigenwerte und Eigenfunktionen verwendeten Begriffsbildungen sind bei dieser Zuordnung invariant; also müssen dieselben in den in I. betrachteten kontinuierlichen Räumen und im diskreten Raume aus II. dasselbe Verhalten zeigen.

---

[7]) Die vollstetigen und die beschränkten Formen (bzw. Operatoren) wurden von Hilbert entdeckt, und er löste ihr Eigenwertproblem (Göttinger Nachr. 1906, S. 157—209), weiter untersucht wurden die beschränkten Bilinearformen von Hellinger, vgl. Anm. [4]), sowie Journal f. Math. **136** (1909), S. 210—273.

[8]) Z. B. Göttinger Nachr. 1907, S. 116—122.

[9]) Den Beweis dieser, allgemein bekannten, Tatsachen (zusammen mit dem Fischer-F. Rieszschen Satze) werden wir im Kap. I und im Anhang I erbringen.

Diese Überlegungen sind es bekanntlich im wesentlichen, durch welche Hilbert die Eigenwerttheorie der Integraloperatoren in kontinuierlichen Räumen auf die der H. O. für Folgen $(x_1, x_2, \ldots)$ (d. h. Funktionen des diskreten Raumes $1, 2, \ldots$) zurückgeführt hat[10]).

IV. Der Gegenstand dieser Arbeit ist die Aufstellung einer allgemeinen Eigenwerttheorie *aller* H. O. Dabei werden wir abstrakt vorgehen, d. h. es von vornherein so einrichten, daß sich unsere Resultate gleichmäßig auf alle in I. genannten kontinuierlichen Räume anwenden lassen, sowie auf den diskreten Raum in II. (d. h. auf die Operatoren der Funktionen in denselben).

Zu diesem Zwecke werden wir im Kap. I eine allgemeine Charakterisierung (durch fünf Eigenschaften **A** bis **E**, vgl. dort) angeben, die auf einen jeden der folgenden Funktionenräume paßt:

Alle (komplexen, im Lebesgueschen Sinne meßbaren) Funktionen f in einem Raume Ω (Zahlengerade, Zahlenebene, Oberfläche der Einheitskugel, Intervall $0, 1$ usw., vgl. Anhang I) mit dem Volumelement dv, für die $\int\limits_\Omega |f|^2 \cdot dv$ endlich ist. Alle Folgen $(x_1, x_2, \ldots)$ von (komplexen) Zahlen mit endlichem

$$\sum_{n=1}^{\infty} |x_n|^2.$$

(Zum ersten Falle ist zu bemerken, daß solche Funktionen f, g, für die $f(P) = g(P)$ überall — mit Ausnahme einer Lebesgueschen 0-Menge — gilt, als nicht verschieden anzusehen sind.)

Die Elemente dieser Funktionenräume lassen alle Vektor-Operatoren zu: man kann sie addieren und mit (komplexen) Zahlen multiplizieren, ferner ist für zwei von ihnen das „innere Produkt" $\int\limits_\Omega f \cdot \bar{g} \cdot dv$ bzw. $\sum\limits_{n=1}^{\infty} x_n \overline{y_n}$ definiert. Das innere Produkt bezeichnen wir mit (f, g) (wenn nichts Besonderes gesagt wird, wollen wir auch die Folgen $(x_1, x_2, \ldots), (y_1, y_2, \ldots), \ldots$ als Funktionen ansehen — vgl. II. — und mit $f, g, \ldots$ bezeichnen), mit seiner Hilfe ist in allen Fällen der Hermitesche Charakter der Operatoren R zu definieren:

$$(f, Rg) = (Rf, g).$$

(In I. war das die Definition, und der in II. ist es gleichwertig.) Wie bei Vektoren ermöglicht das innere Produkt die Definition eines Absolutwertes: $|f| = \sqrt{(f, f)} \left(\sqrt{\int\limits_\Omega |f|^2 \cdot dv} \text{ bzw. } \sqrt{\sum\limits_{n=1}^{\infty} |a_n|^2}\right)$, und einer Entfernung:

---

[10]) Dies ist insofern historisch ungenau, als der Fischer-F. Rieszsche Satz erst nachher gefunden wurde. Dieselben Überlegungen haben in der neuen Quantentheorie eine wichtige Rolle gespielt, vgl. Schrödinger, Annalen d. Phys. **79**, 8 (1926), S. 734—756.

$|f-g|$ (diese ist also das Analogon der gewöhnlichen Euklidischen Entfernung); unser Raum wird demnach durch dasselbe „metrisiert“ und „topologisiert“. D. h. topologische Ausdrucksweisen, wie Abgeschlossenheit, Stetigkeit usw., werden in ihm sinnvoll.

Es ist vielleicht nicht überflüssig, noch ein Wort der konsequenten Beschränkung auf Funktionen f mit endlichem $\int\limits_{\Omega} |f|^2 \cdot dv$ zu widmen. Für Eigenwertprobleme bei Folgen ist es seit den grundlegenden Untersuchungen von Hilbert (vgl. Anm. [7]) und E. Schmidt[11]) üblich, die Endlichkeit von $\sum\limits_{n=1}^{\infty} |x_n|^2$ zu verlangen; bei den f in kontinuierlichen Räumen ist jedenfalls für die Eigenfunktionen des Punktspektrums $\int\limits_{\Omega} |f|^2 \cdot dv$ endlich. Im Streckenspektrum ist freilich weder $\sum\limits_{n=1}^{\infty} |x_n|^2$ noch $\int\limits_{\Omega} |f|^2 \cdot dv$ endlich (wenn überhaupt noch Eigenfolgen bzw. -funktionen existieren und nicht zu „differentiellen Lösungen“ Zuflucht genommen werden muß), indessen werden wir (unter Verallgemeinerung der Hilbertschen Methoden) zeigen, daß es gerade günstig ist, die Eigenfunktionen im Streckenspektrum als uneigentliche Gebilde zu behandeln, d. h. das Eigenwertproblem ohne ihre explizite Nennung zu formulieren.

Unser Funktionenraum ist, wenn wir den diskreten Raum $1, 2, \ldots$ zugrunde legen, der sogenannte (komplexe) Hilbertsche (d. h. unendlichvieldimensionale Euklidische) Raum: die Menge aller Folgen $(x_1, x_2, \ldots)$ mit endlichem $\sum\limits_{n=1}^{\infty} |x_n|^2$. Da, wie in III. bemerkt wurde, alle unsere Funktionen-Mannigfaltigkeiten isomorph sind (d. h. ein-eindeutig aufeinander abgebildet werden können, wobei die Operationen $f+g$, $a \cdot f$, (f, g) in ihre Analoga übergehen), wollen wir sie auch als Hilbertsche Räume bezeichnen. Und sie alle sollen als spezielle Verwirklichungen des allgemeinen abstrakten Hilbertschen Raumes angesehen werden, der durch die „inneren“ Eigenschaften **A** bis **E** allein (und bis auf Isomorphien vollständig, vgl. Kap. I) charakterisiert ist.

Den abstrakten Hilbertschen Raum nennen wir $\mathfrak{H}$.

V. Im abstrakten Hilbertschen Raume verstehen wir unter einem H. O., wie in I. gesagt wurde, eine Funktion R (für gewisse f aus $\mathfrak{H}$ definiert und Werte aus $\mathfrak{H}$ annehmend), die die folgenden Eigenschaften besitzt:

$$(\alpha) \qquad R(a_1 f_1 + \ldots + a_k f_k) = a_1 R f_1 + \ldots + a_k R f_k$$
$$(a_1, \ldots, a_k \text{ komplexe Konstanten}),$$

$$(\beta) \qquad (f, Rg) = (Rf, g)$$

---

[11]) Z. B. Rend. d. Circ. Mat. d. Palermo **25** (1908), S. 57–73.

(soweit beide Seiten Sinn haben). Wir nehmen übrigens einstweilen an, daß der Definitionsbereich von R eine lineare Mannigfaltigkeit ist, d. h. mit $f_1, \dots, f_k$ auch $a_1 f_1 + \dots + a_k f_k$ enthält. (Die endgültige Definition der H. O. wird zwar etwas anders lauten, es kommt uns aber zunächst darauf nicht an.)

Wir nennen einen H. O. R beschränkt, wenn $|f| \leqq 1$ $|Rf| \leqq C$ zur Folge hat, wobei C beliebig, aber von f unabhängig ist. Da allgemein $|af| = |a|\,|f|$ gilt, ist dies mit $|Rf| \leqq C \cdot |f|$, oder auch $|Rf - Rg| \leqq C \cdot |f - g|$ gleichbedeutend: d. h. es drückt eine Art Stetigkeit von R aus[12]).

In der von Hilbert aufgebauten Theorie der beschränkten Hermiteschen Formen (d. h. H. O.) kann man stets annehmen, daß R für alle f von $\mathfrak{H}$ definiert ist: wenn das nicht der Fall ist, kann es (wegen seiner Stetigkeit) auf ganz $\mathfrak{H}$ fortgesetzt werden. Wenn wir aber eine allgemeine Theorie der H. O. aufstellen wollen, die auch über die beschränkten hinausgeht, so dürfen wir nicht verlangen, daß R überall in $\mathfrak{H}$ definiert sei. (Nach einem Satze von Toeplitz muß sogar jeder überall sinnvolle H. O. beschränkt sein, vgl. Satz 48, α), γ) und Anm. [61]).) So sei z. B. Ω das Intervall 0, 1, und R der Operator $i\frac{d}{dx}\dots$ für alle differentiierbaren, in 0 und 1 verschwindenden Funktionen. R ist ein H. O., aber es ist offenbar nicht für undifferentiierbare Funktionen zu definieren! Ebensowenig wird für einen hochsingulären Integralkern $\varphi(P, P')$ $\int\limits_{\Omega} \varphi(P', P) \cdot f(P') \cdot dv'$ für alle f (mit endlichem $\int\limits_{\Omega} |f|^2 \cdot dv$) konvergieren. Ein anders geartetes, aber auch recht charakteristisches Beispiel ist dieses: Ω sei die Zahlengerade, $Rf(x) = x \cdot f(x)$. Auch R ist ein H. O., da aber $\int\limits_{-\infty}^{\infty} |f(x)|^2 dx$ endlich sein kann bei unendlichem $\int\limits_{-\infty}^{\infty} x^2 |f(x)|^2 dx$, ist auch dieses R nicht stets sinnvoll.

Die Einschränkung also, daß (α), (β) nur soweit zu gelten haben, als beide Seiten sinnvoll sind, ist eine recht wesentliche. Immerhin werden wir aber (als Minimum) verlangen, daß

(γ) der Definitionsbereich von R überall dicht (in $\mathfrak{H}$) sei.

Was dies bedeutet, kann man sich z. B. an $i\frac{d}{dx}\dots$ bzw. $x \cdot \dots$ mühelos

---

[12]) Wenn wir $\mathfrak{H}$ als Folgenraum (eigentlichen Hilbertschen Raum) auffassen, so ist unsere Definition der Beschränktheit dem Wortlaute nach von der Hilbertschen verschieden: wir verlangen (in Hilberts Terminologie) die Beschränktheit des mit sich selbst gefalteten R. Immerhin sind diese Formulierungen gleichwertig, näher erörtern wir dies in Kap. II, Satz 12. — Man beachte übrigens, daß unsere ganze Topologie von $\mathfrak{H}$ stets im Sinne der sogenannten „starken Konvergenz" zu verstehen ist, vgl. Weyl, Dissertation, S. 9.

klarmachen: diese sind für alle (in 0 und 1 verschwindenden) Polynome bzw. für alle außerhalb eines willkürlichen (aber endlichen) Intervalles stets verschwindenden Funktionen immer sinnvoll, und beide Funktionen-Mengen sind überall dicht.

Bei dieser Anordnung laufen wir aber Gefahr, zu wenig zu verlangen: unsere H.O. könnten bereits an solchen Stellen (infolge einer willkürlichen Einengung des Definitionsbereiches) sinnlos sein, wo sie noch auf natürliche Weise definiert werden könnten. (So hätte man z. B. bei $i\frac{d}{dx}\cdots$, unbeschadet der überall-Dichtigkeit des Definitionsbereiches, statt einmaliger Differentiierbarkeit auch Analytizität verlangen können.) Um dieser Schwierigkeit zu entgehen, stellen wir den Begriff des maximalen H.O. auf:

Der H.O. S heiße Fortsetzung des H.O. R, wenn überall wo Rf Sinn hat, auch Sf Sinn hat, und mit Rf übereinstimmt. Wenn $S \neq R$ ist (d. h. S an mehr Stellen sinnvoll ist, wie R), so ist S eine eigentliche Fortsetzung. Ein H.O. R, zu dem es keine eigentlichen Fortsetzungen mehr gibt, heiße maximal. (Abkürzungen: Forts., max.)

Da, wie wir sehen werden (Satz 35), jeder H.O. zu einem max. H.O. fortsetzbar ist, müssen wir in erster Linie die max. H.O. untersuchen. Dazu ist es aber notwendig, eine allgemeine Formulierung des Eigenwertproblems anzugeben (und zwar im abstrakten Hilbertschen Raume) bei der Punkt- und Streckenspektrum gleichmäßig Berücksichtigung finden. Dies soll im nächsten Paragraphen, in Anlehnung an Hilberts Formulierung des Problems, erfolgen.

VI. Betrachten wir zunächst den k-dimensionalen Euklidischen Raum. Wenn dort ein H.O. gegeben ist, d. h. eine Zuordnung

$$R(x_1, \ldots, x_k) = (y_1, \ldots, y_k),$$

$$y_n = \sum_{m=1}^{k} a_{mn} x_m \quad (n = 1, \ldots, k), \qquad a_{mn} = \overline{a_{nm}}$$

(Konvergenzfragen gibt es hier ja nicht!), so ist es bekanntlich stets möglich R durch Einführung neuer kartesischer Koordinaten (d. h. durch eine unitäre Transformation) auf die Diagonalform zu bringen. D. h. es gibt eine unitäre Transformationsmatrix $\{u_{mn}\}$

$$\sum_{p=1}^{k} u_{mp}\overline{u_{np}} = \left\{\begin{matrix} 1, & \text{für} & m = n \\ 0, & \text{für} & m \neq n \end{matrix}\right|, \qquad \sum_{p=1}^{k} u_{pm}\overline{u_{pn}} = \left\{\begin{matrix} 1, & \text{für} & m = n \\ 0, & \text{für} & m \neq n \end{matrix}\right|,$$

so daß nach der Transformation

$$\xi_n = \sum_{m=1}^{k} u_{mn} x_m, \qquad \eta_n = \sum_{m=1}^{k} u_{mn} y_m \qquad (n = 1, \ldots, k)$$

R einfach durch

$$\eta_n = \lambda_n \xi_n \qquad (n = 1, \ldots, k)$$

($\lambda_1, \ldots, \lambda_k$ reelle Konstante, die Eigenwerte von R!) beschrieben wird[13]). D. h. es ist

$$a_{mn} = \sum_{p=1}^{k} \lambda_p u_{mp} \overline{u_{np}} \qquad (m, n = 1, \ldots, k).$$

Dabei sind die $\lambda_1, \ldots, \lambda_k$, bis auf die Reihenfolge, durch R eindeutig bestimmt, nicht aber die u_{mn}! Wie man sofort sieht, kann man sie durch $\vartheta_n u_{mn}$ ersetzen ($\vartheta_1, \ldots, \vartheta_k$ beliebige Zahlen vom Absolutwerte 1), und wenn mehrere der λ_n einander gleich sind, sind sogar die dazugehörigen Zeilen ($u_{1n}, \ldots, u_{kn}$) beliebig untereinander unitär transformierbar.

Eine uneindeutige Formulierung des Eigenwertproblems, wie diese (Transformierbarkeit auf die Diagonalform), ist wenig geeignet, um von den k-dimensionalen Euklidischen Räumen auf den unendlichvieldimensionalen (d. h. Hilbertschen) übertragen zu werden. Die richtige, leicht zu verallgemeinernde, Fassung des endlichvieldimensionalen Problems findet man, indem man es nach Hilbert eindeutig formuliert. Dies geschieht so:

Wir setzen

$$e_{mn}(\lambda) = \sum_{\lambda_p \leq \lambda} u_{mp} \overline{u_{np}} \qquad (\lambda \text{ eine beliebige reelle Zahl}),$$

dann sind die $\{e_{mn}(\lambda)\}$ Matrizen von H. O. $E(\lambda)$, mit der leicht zu verifizierenden Eigenschaft

$$(\alpha) \qquad E(\lambda)E(\mu) = E(\mu)E(\lambda) = E(\lambda) \qquad (\lambda \leq \mu)$$

(also insbesondere $E(\lambda)^2 = E(\lambda)$)[14]). Seien $l_1, \ldots, l_h$ ($h \leq k$) die voneinander verschiedenen Werte der $\lambda_1, \ldots, \lambda_k$, und zwar wachsend angeordnet. Dann ist $E(\lambda)$ (als Funktion von λ) in den Intervallen

$$\lambda < l_1,\ l_1 \leq \lambda < l_2,\ \ldots,\ l_{h-1} \leq \lambda < l_h,\ l_h \leq \lambda$$

konstant, und bei $l_1, \ldots, l_h$ liegen (links) Sprünge. Im ersten der genannten Intervalle ist es offenbar $= 0$, im letzten (wegen der Unitarität von $\{u_{mn}\}$) hat es die Einheitsmatrix, es ist also der identische Operator 1. Aus der letzten Formel für a_{mn} folgt sofort (l_0 beliebig, aber $< l_1$):

$$a_{mn} = \sum_{p=1}^{h} l_p (e_{mn}(l_p) - e_{mn}(l_{p-1})) = \int_{-\infty}^{\infty} \lambda \, d\, e_{mn}(\lambda).^{15})$$

---

[13]) Wie man sieht, sind $\lambda_1, \ldots, \lambda_k$ Eigenwerte, und $(\overline{u_{11}}, \ldots, \overline{u_{k1}}), \ldots, (\overline{u_{1k}}, \ldots, \overline{u_{kk}})$ dazugehörige Eigenvektoren (vgl. II.).

[14]) Was bei überall sinnvollen Operatoren A, B unter AB zu verstehen ist, ist klar.

[15]) Stieljessches Integral!

Hieraus folgt sofort (wenn, ebenso wie in IV., $(f, g) = \sum_{n=1}^{k} x_n \overline{y_n}$ ist, falls $f = (x_1, \ldots, x_k)$, $g = (y_1, \ldots, y_k)$ ist)

$$(\beta) \qquad (f, Rg) = \int_{-\infty}^{\infty} \lambda\, d(f, E(\lambda) g).$$

Man überlegt sich leicht, daß $(\alpha), (\beta)$, zusammen mit $E(\lambda) = 0$ bzw. 1 für kleine bzw. große λ, und damit, daß $E(\lambda)$ in λ nach rechts stetig ist, die Operatoren $E(\lambda)$ vollkommen festlegt. Da aber aus den $E(\lambda)$ (bzw. den $e_{mn}(\lambda)$) die Diagonalform von R leicht erzeugt werden kann, ist dies die eindeutige Formulierung des Eigenwertproblems.

Kehren wir nun zum abstrakten Hilbertschen Raume $\mathfrak{H}$ zurück! Die Bedingungen $(\alpha), (\beta)$ dürfen da wörtlich übertragen werden, wenn wir nur die $E(\lambda)$ genau erklären: sie sollen überall sinnvolle H. O. sein[16]). Die Zusatzforderungen betreffs der λ-Abhängigkeit der $E(\lambda)$ formulieren wir so:

$$(\gamma) \qquad \text{für } \lambda \to -\infty \quad E(\lambda) \to 0, \qquad \text{für } \lambda \to +\infty \quad E(\lambda) \to 1,$$
$$\text{für } \lambda \geqq \lambda_0,\ \lambda \to \lambda_0 \quad E(\lambda) \to E(\lambda_0).$$ [17])

(Daß gerade dies die richtige Generalisierung ist, wird der Erfolg zeigen.)

Eine $(\alpha), (\gamma)$ genügende Schar $E(\lambda)$ heiße eine Zerlegung der Einheit (kurz: Z. d. E.); wenn R zu ihr in der Beziehung (β) steht, ist es der dazugehörige Operator. Freilich entstehen Konvergenzfragen, wir werden aber sehen (vgl. Satz 36), daß diese am besten behoben werden, wenn wir (β) so verschärfen: Rg habe dann und nur dann Sinn, wenn

$$(\beta') \qquad \int_{-\infty}^{\infty} \lambda^2\, d\,|E(\lambda) g|^2$$

endlich ist — dann sind (wie wir zeigen werden) alle Integrale von (β) absolut konvergent — und es möge (β) gelten. (Aus (α) folgt, daß $|E(\lambda) f|^2$ als Funktion von λ nie abnimmt, vgl. Satz 17, 18, der Integrand λ^2 von (β') ist $\geqq 0$: also ist das Integral (β') entweder absolut konvergent, oder eigentlich divergent, und zwar $+\infty$.) Man sieht leicht ein: eine Z. d. E. (also eine die $(\alpha), (\gamma)$ genügt) bestimmt den zugehörigen Operator (wenn es einen gibt) eindeutig: denn der Definitionsbereich ist gegeben (durch (β')), und alle (f, Rg) (durch (β)), also auch Rg.

---

[16]) Im k-dimensionalen Raume hat $E(\lambda)^2 = E(\lambda)$, wie man leicht verifiziert, die Konsequenz $|E(\lambda) f| \leqq |f|$; es ist also zu erwarten, daß es in $\mathfrak{H}$ ebenso sein wird, also daß $E(\lambda)$ als beschränkt und somit überall sinnvoll anzusetzen ist (vgl. auch Kap. III, Satz 14, 15, 16).

[17]) Bei (überall sinnvollen) Operatoren A_n, B verstehen wir unter $A_n \to B$, daß für alle f $A_n f \to Bf$ gilt. 1 ist durch $1f = f$ definiert.

Das Eigenwertproblem besteht also darin, zu entscheiden: Gibt es zu jeder Z. d. E. einen zugehörigen H. O.? Ist dieser maximal? Müssen zu verschiedenen Z. d. E. verschiedene Operatoren gehören? Welche H. O. gehören zu Z. d. E.?[18]) Wir werden sehen, daß die drei ersten Fragen bejahend zu beantworten sind (vgl. Satz 36). Bei der vierten müssen wir daher weiter fragen: Die den Z. d. E. zugeordneten H. O. sind eine Teilklasse der Klasse aller max. H. O. Welche ist nun diese Teilklasse?

VII. Wir werden sehen (Satz 36, 38), daß sie nicht alle max. H. O. umfaßt. Es gibt eine Ausnahme[19]) $\overline{R}$, und alle anderen Ausnahmen lassen sich durch gewisse, hier nicht zu erörternde, triviale Operationen aus ihr und den H. O. mit Z. d. E. aufbauen (vgl. Satz 38). Bei $\overline{R}$ werden wir sehen, daß seine Eigenschaften in der Tat jede vernünftige Formulierung eines Eigenwertproblems ausschließen, daß es aber keineswegs „pathologisch" ist: es ist leicht zu beschreiben (vgl. Anm.[19])).

Immerhin müssen solche max. H. O., die beschränkt, oder definit, oder reell[20]) sind, eine Z. d. E. besitzen, gewisse einfache Eigenschaften sind somit mit dem Fehlen der Z. d. E. unverträglich (vgl. Satz 40). Das Auftreten dieses $\overline{R}$ können wir auch so deuten, daß der Hermitesche Charakter bei unendlichen Operatoren das Auftreten der Ausnahmefälle der Elementarteilertheorie nicht mehr verhindert (im Gegensatze zum Falle endlich vieler Dimensionen; oder im Unendlichvieldimensionalen, zu den reellen, oder beschränkten, oder definiten Operatoren), aber daß nur ein einziger Ausnahme-Typus vorkommen kann.

Es genügt wohl vorläufig soviel über die max. H. O. zu sagen. (Insbesondere soll nicht weiter erörtert werden, wie und in welchem Umfange die Kenntnis der Z. d. E. — d. h. der „Hilbertschen Spektralform" — die Lösung des Eigenwertproblems, so wie es in I., II. formuliert wurde, ermöglicht. Es sei diesbezüglich auf die in Anm. [4]), [5]) zitierten Untersuchungen

---

[18]) Hilbert bewies: Für beschränkte H. O. gibt es stets eine und nur eine Z. d. E., und zwar kommen die und nur die Z. d. E. vor, für die $E(\lambda)$ für hinreichend kleine wie für hinreichend große λ konstant (also $= 0$ bzw. 1) wird. (β') fällt dann fort, da es stets erfüllt ist. (Zitat: Anm. [7]).) — Bei unserer Formulierung ist offenbar die Abtrennung von Punkt- und Streckenspektrum noch nicht vollzogen (in (γ) wird $\lambda \geqq \lambda_0$ verlangt!), sie ist für unsere Zwecke unerheblich.

[19]) Der unter anderem der Anhang IV dieser Arbeit gewidmet wird.

[20]) Sei $\varphi_1, \varphi_2, \ldots$ ein vollständiges Orthogonalsystem in $\mathfrak{H}$ (vgl. Def. 3), dann kann jedes f eindeutig auf die Form $\sum_{n=1}^{\infty} x_n \varphi_n$ gebracht werden, es heiße reell, wenn alle x_n es sind. $\Re f$ entsteht aus f, wenn alle x_n durch $\Re x_n$ (Realteil!) ersetzt werden. R ist reell, wenn bei sinnvollem Rf auch $R \, \Re f$ sinnvoll ist, und zwar $= \Re Rf$. (In den Begriff der „Realität" geht also noch willkürlich ein gewisses vollständiges Orthogonalsystem ein.)

verwiesen.) Dagegen ist noch der Fall solcher H. O. zu erörtern, deren Maximalität nicht gesichert (oder nicht vorhanden) ist. Die H. O. sind ja meistens so gegeben: z. B. als selbstadjungierte Differentialoperatoren, definiert für alle n-mal differentiierbaren Funktionen (n ist der Grad!), die gewissen Randbedingungen genügen. Wir haben in V. erwähnt, daß jeder H. O. zu einem max. H. O. fortsetzbar ist, es ist aber nicht gesagt, daß dies nur auf eine einzige Weise geht.

Wir werden auch diesbezüglich einfache Resultate erzielen, so u. a.: wenn ein H. O. nicht selbst max. oder nicht nach einer gewissen trivialen Forts. max. ist, so kann er auf Kontinuum viele verschiedene Arten max. gemacht werden (vgl. die Schlußbemerkung von Kap. VIII). Beschränkte H. O. sind stets im ersten Falle.

Es werde noch an einem einfachen Beispiel klargemacht, wie diese mehrdeutige Fortsetzbarkeit zustande kommt. Sei Ω das Intervall 0, 1, R der Operator $i\frac{d}{dx}\ldots$, definiert für alle differentiierbaren Funktionen. Da

$$(f, Rg) - (Rf, g) = \int_0^1 i\{f(x)\bar{g}'(x) + f'(x)\bar{g}(x)\}\,dx$$
$$= i\{f(1)\bar{g}(1) - f(0)\bar{g}(0)\}$$

ist, ist R nicht Hermitesch, wohl aber R_0 und R_ϑ (ϑ eine Zahl vom Absolutwerte 1), die aus R durch die folgende Einschränkung des Definitionsbereiches entstehen:

$$f(0) = f(1) = 0 \quad \text{bzw.} \quad f(1) = \vartheta \cdot f(0).$$

R_ϑ hat max. Fortsetzungen, z. B. S_ϑ. Alle S_ϑ sind auch max. Forts. von R_0. Und sie sind alle verschieden, denn wäre $S_{\vartheta_1} = S_{\vartheta_2} = S$, so wäre S für ein f und ein g mit $f(0) = g(0) = 1$, $f(1) = \vartheta_1$, $g(1) = \vartheta_2$ sinnvoll, und daher

$$(f, Sg) - (Sf, g) = \vartheta_1\bar{\vartheta}_2 - 1 \neq 0,$$

also S kein H. O. Die Mehrdeutigkeit der max. Forts. rührt hier also von der Mehrzahl der möglichen Randbedingungen her. Wir werden sehen, daß noch im allgemeinsten Falle die Verhältnisse eine gewisse Analogie hiermit aufweisen (vgl. Satz 26, 27)[21]).

VIII. Der Zusammenhang der H. O. von $\mathfrak{H}$ mit den Matrizen ist bekanntlich dieser: sei R ein H. O. und $\varphi_1, \varphi_2, \ldots$ ein vollständiges Orthogonalsystem (vgl. Def. 3), dann wird R die Matrix $\{a_{mn}\}$, $a_{mn} = (\varphi_m, R\varphi_n)$, zugeordnet. Diese Zuordnung ist bei beschränkten H. O. ganz klar, bei

---

[21]) Bei singulären Integraloperatoren fand schon Carleman Zeichen dieser Mehrdeutigkeit, jedoch konnte er in diesem Falle das Eigenwertproblem nicht ganz lösen (es fehlte das Analogon der entscheidenden Bedingung (α) aus IV.). Vgl. Anm. [4]).

unbeschränkten ist sie mit gewissen Komplikationen verknüpft. Sie wird uns im Anhang III beschäftigen. Die unitäre Transformationstheorie der unbeschränkten Matrizen führt in eine besonders eigentümliche Pathologie, dieselbe soll aber an anderer Stelle untersucht werden[22]).

Auch über das Verhältnis dieser Resultate zur neuen Quantentheorie (der sogenannten „Transformationstheorie") wollen wir hier nicht sprechen, jedenfalls zeigen sie, daß die allgemeine Theorie der H. O. keineswegs überall dasjenige Verhalten zeigt, welches in der „Transformationstheorie" (auf Grund der Analogie mit beschränkten oder gar endlichvieldimensionalen Operatoren) im allgemeinen vermutet und vorweggenommen wurde.

IX. Es bleibt noch übrig, einiges über die Methode und Anordnung dieser Arbeit zu sagen.

Der wesentliche Kunstgriff ist das Untersuchen eines gewissen Operators U, der symbolisch mit $U = \frac{R + i1}{R - i1}$ bezeichnet werden könnte, an Stelle von R. Wenn R ein H. O. ist, so ist nämlich U längentreu (d. h. $|Uf| = |f|$), also stetig, und somit leichter zu behandeln (vgl. Satz 22, 23, 24). Diese „Cayley-Transformation" $U = \frac{R + i1}{R - i1}$ ist der Kern unserer Methode.

Alle anderen Methoden versagen: Die eleganten Verfahren von Hellinger und von F. Riesz, weil in ihnen der Operator R iteriert werden muß, was nur bei überall sinnvollen (also beschränkten) H. O. ohne weiteres geht, bei beliebigen H. O. aber zunächst fraglich ist. Alle Maximum-Minimum-Methoden, sowie die Methode von E. Schmidt sind von vornherein auf Fälle ohne Streckenspektrum beschränkt. Die ursprüngliche Hilbertsche Methode (Approximation durch endlichvieldimensionale Abschnittoperatoren) ist es allein, die gewisse Resultate zu erzielen erlaubt: aber nur unter sehr großen Schwierigkeiten, und nur einen Bruchteil dessen, was wir erreichen können[23]).

---

[22]) Vgl. eine demnächst im Journal f. Math. erscheinende Arbeit des Verfassers: „Zur Theorie der unbeschränkten Matrizen".

[23]) Der Verfasser konnte mit dieser Methode den Fall reeller H. O. (vgl. Anm. [20])) erledigen, wo der in VII. angedeutete Ausnahmefall noch nicht in Erscheinung trat. Dabei konnte aber der Fortsetzungsprozeß der H. O. nicht so übersehen werden, wie es etwa im Kap. VIII der Fall sein wird, und die Methode war derart unkonstruktiv, daß dabei z. B. der Wohlordnungssatz herangezogen werden mußte (vgl. die Ankündigung im Jahresber. d. D. Math.-Ver. **37**, 1–4 (1928), S. 11–15), auch war den nicht-reellen H. O. so nicht beizukommen. — Der Verfasser möchte es nicht versäumen, Herrn E. Schmidt an dieser Stelle seinen Dank für das Interesse auszusprechen, das er diesen Untersuchungen (insbesondere der Übertragung der Resultate auf nicht-reelle H. O.) zugewandt hat, und darauf hinweisen, daß er wiederholt wesentliche Anregungen aus Gesprächen mit ihm empfing. Insbesondere stammt der Begriff der „Hypermaximalität" von H. O. (vgl. Def. 9), sowie die Erkenntnis, daß diese Eigenschaft für die Existenz der Z. d. E. notwendig und hinreichend ist, von E. Schmidt.

Die Anordnung des Gegenstandes ist die folgende: Die Kap. I bis IV sind vorbereitender Natur: I. Struktur des Raumes $\mathfrak{H}$, II. Allgemeines über Operatoren in $\mathfrak{H}$, III. Lineare Mannigfaltigkeiten und ihre Projektionsoperatoren, IV. Reduzibilität von Operatoren. In den Kap. V bis VIII wird die Hauptarbeit der Lösung des Eigenwertproblems geleistet: V. Cayleysche Transformation, VI. Erweiterungselemente, VII. Defektindizes, VIII. Der Fortsetzungsprozeß. In Kap. IX bis X werden im Anschluß daran die genauen Normalformen für die max. H. O. hergestellt: IX. Eigenwertdarstellung, X. Operatoren ohne Eigenwertdarstellung. Kap. XI bis XII enthalten schließlich einige allgemeine Sätze über unbeschränkte Operatoren: XI. Halbbeschränktheit, XII. Pathologie der Unbeschränktheit[24]). Außerdem sind noch einige Anhänge da: In I. wird bewiesen, daß die Funktionenräume von I. (Einleitung) wirklich den Bedingungen **A** bis **E** von Kap. I genügen (also daß unsere Theorie auf sie anwendbar ist)[25]). In Anhang II beweisen wir einen Satz über die Lösung des Eigenwertproblems unitärer Operatoren (allgemeiner dessen von „normalen") — dies ist eine Frage über beschränkte Operatoren, die mit der F. Rieszschen Methode lösbar ist, aber im Rahmen der Hilbertschen Theorie bisher nie erledigt wurde[26]). Wir brauchen dies für Kap. VIII, da es aber methodisch in den Kreis der Hilbertschen Theorie gehört, ziehen wir vor, es so getrennt zu behandeln. In Anhang III gehen wir auf den Zusammenhang von Operatoren und Matrizen näher ein (vgl. VIII. Einleitung, sowie Anm.[23])); in IV wird der Operator $\bar{R}$ (vgl. VII. Einleitung) genauer untersucht.

Im übrigen ist diese Arbeit ein logisch geschlossenes Ganzes, das ohne alle Vorkenntnisse aus der Literatur der „unendlich vielen Variablen" gelesen werden kann. (Dementsprechend ist auch der Inhalt der Kap. I, III und des Anhangs I — abgesehen von der Anordnung — nicht neu.)

I. Der abstrakte Hilbertsche Raum.

Wir charakterisieren den abstrakten (komplexen) Hilbertschen Raum durch fünf Eigenschaften **A** bis **E**. Es erweist sich als zweckmäßig, an einige von ihnen einige einfache Sätze anzuschließen, die bereits aus den betreffenden allein folgen.

---

[24]) Dieselbe wird in der in Anm.[22]) angekündigten Arbeit des Verfassers noch bedeutend verschärft werden.

[25]) Dies ergibt, zusammen mit einigen Überlegungen von Kap. I, natürlich einen Beweis des Fischer-F. Rieszschen Satzes.

[26]) Es kommt im wesentlichen darauf heraus, daß zwei vertauschbare beschränkte H. O. eine gemeinsame Eigenwertform (oder Spektralform) nach Hilbert zulassen. Vgl. Hellinger-Toeplitz, Encyklopädie-Artikel II, C. 13, S. 1562 u. f., wo die Frage für vollstetige H. O. gelöst wird.

A. *$\mathfrak{H}$ ist ein linearer Raum.*

D. h.: es gibt in $\mathfrak{H}$ eine Addition $f+g$, eine Subtraktion $f-g$, und eine (skalare) Multiplikation af (f, g von $\mathfrak{H}$, a eine komplexe Zahl), sowie ein Element 0; und für diese gelten die Rechenregeln der gewöhnlichen Vektoralgebra.

Definition 1. Eine Teilmenge $\mathfrak{M}$ von $\mathfrak{H}$ heißt eine lineare Mannigfaltigkeit (kurz: lin. M.), wenn sie mit $f_1, \ldots, f_n$ auch $a_1 f_1 + \ldots + a_n f_n$ enthält. Wenn $\mathfrak{M}$ irgendeine Teilmenge von $\mathfrak{H}$ ist, so ist die Menge aller $a_1 f_1 + \ldots + a_n f_n$ ($f_1, \ldots, f_n$ von $\mathfrak{M}$, $a_1, \ldots, a_n$ Konstante) die kleinste $\mathfrak{M}$ enthaltende lin. M., sie heiße die von $\mathfrak{M}$ aufgespannte lin. M.

Definition 2. n Elemente $f_1, \ldots, f_n$ heißen linear unabhängig (kurz: lin. unabh.), wenn aus $a_1 f_1 + \ldots + a_n f_n = 0$ $a_1 = \ldots = a_n = 0$ folgt. n lin. M. $\mathfrak{M}_1, \ldots, \mathfrak{M}_n$ heißen lin. unabh., wenn aus $f_1 + \ldots + f_n = 0$ (f_1 von $\mathfrak{M}_1, \ldots, f_n$ von $\mathfrak{M}_n$) $f_1 = \ldots = f_n = 0$ folgt.

B. *Es gibt in $\mathfrak{H}$ ein, zu dem der Vektorrechnung analoges, inneres Produkt, das eine Metrik erzeugt.*

D. h.: es gibt eine Funktion (f, g) (f, g von $\mathfrak{H}$, (f, g) eine komplexe Zahl) mit den folgenden Eigenschaften:

1. $(af, g) = a(f, g)$, 2. $(f_1 + f_2, g) = (f_1, g) + (f_2, g)$, 3. $(f, g) = \overline{(g, f)}$,

4. $(f, f) \geqq 0$ *und nur für* $f = 0$ *ist es* $= 0$.

(*Aus* 1., 2. *folgt nach* 3.: 1'. $(f, ag) = \bar{a}(f, g)$, 2'. $(f, g_1 + g_2) = (f, g_1) + (f, g_2)$.) *Durch*

$$|f| = \sqrt{(f, f)}$$

wird ein „Betrag" definiert, durch $|f - g|$ *die Metrik* (*vgl. Satz* 2 *und Anm.*[27])).

Satz 1. *Es gilt stets*

$$|(f, g)| \leqq |f| \cdot |g|.$$

Beweis. Erstens ist (nach 2., 2'., 3. und 4.)

$$\Re(f, g) = \left(\frac{f+g}{2}, \frac{f+g}{2}\right) - \left(\frac{f-g}{2}, \frac{f-g}{2}\right) \leqq \left(\frac{f+g}{2}, \frac{f+g}{2}\right) + \left(\frac{f-g}{2}, \frac{f-g}{2}\right)$$
$$= \frac{1}{2}((f, f) + (g, g)) = \frac{|f|^2 + |g|^2}{2}.$$

Wir ersetzen f, g durch af, $\frac{1}{a} g$ ($a > 0$): die linke Seite bleibt invariant, die rechte hat das Minimum $|f| \cdot |g|$. Wir ersetzen f durch $\vartheta \cdot f$ ($|\vartheta| = 1$), die rechte Seite bleibt invariant, die linke hat das Maximum $|(f, g)|$. Also

$$|(f, g)| \leqq |f| \cdot |g|.$$

Satz 2. *Es gilt stets*:

$$|af| = |a|\,|f|, \qquad |f+g| \leq |f| + |g|.$$

Beweis. Das erste folgt sofort aus 1., 1'., das zweite aus Satz 1:

$$(f+g, f+g) = (f, f) + (g, g) + 2\,\mathfrak{R}(f, g),$$

$$|f+g|^2 = |f|^2 + |g|^2 + 2\,\mathfrak{R}(f, g) \leq |f|^2 + |g|^2 + 2\,|f|\,|g|,$$

$$|f+g| \leq |f| + |g|.$$

Definition 3. Zwei f, g sind orthogonal (kurz: orth.), wenn $(f, g) = 0$ ist; zwei lin. M. sind orth., wenn jedes Element der einen zu jedem der anderen orth. ist. Eine Menge $\mathfrak{M}$ heißt normiert orthogonal (kurz: norm. orth.), wenn

$$(f, g) = \begin{cases} 1, & \text{für } f = g \\ 0, & \text{für } f \neq g \end{cases} \Bigg| \quad f, g \text{ von } \mathfrak{M}$$

gilt. Sie ist vollständig (kurz: vollst. norm. orth.), wenn sie nicht echter Teil einer anderen norm. orth. Menge ist — was offenbar besagt, daß $(f, g) = 0$ (f fest, für alle g von $\mathfrak{M}$) $|f| = 1$ ausschließt, also auch $|f| > 0$, also $f = 0$ zur Folge hat. (Bekanntlich hat bei mehreren Elementen oder lin. M. die paarweise Orth. die lin. Unabh. zur Folge.)

Definition 4. Nach Satz 2 ist $|f-g|$ eine Entfernung in $\mathfrak{H}$[27]), also $\mathfrak{H}$ topologisiert. Eine abgeschlossene (kurz: abg.) lin. M. ist also eine, die ihre Häufungspunkte enthält. Wenn wir zur von $\mathfrak{M}$ aufgespannten lin. M. alle ihre Häufungspunkte hinzufügen, erhalten wir die kleinste $\mathfrak{M}$ enthaltende abg. lin. M.: die von $\mathfrak{M}$ aufgespannte abg. lin. M.

C. *In der Metrik $|f-g|$ ist $\mathfrak{H}$ separabel. D. h.: eine gewisse abzählbare Menge ist in $\mathfrak{H}$ überall dicht.*

D. *$\mathfrak{H}$ besitzt beliebig (endlich!) viele lin. unabh. Elemente*[28]).

[27]) Von einer „Entfernung" $\overline{f, g}$ verlangt man bekanntlich, daß sie den folgenden Axiomen genüge:

1. $\overline{f, g} \geq 0$, und nur für $f = g$ ist es $= 0$;
2. $\overline{f, g} = \overline{g, f}$;
3. $\overline{f, h} \leq \overline{f, g} + \overline{g, h}$.

In linearen Räumen kommt noch

4. $\overline{f+h, g+h} = \overline{f, g}, \quad \overline{af, ag} = |a| \cdot \overline{f, g}$

hinzu. — Alle diese Eigenschaften besitzt $\overline{f, g} = |f-g|$ nach Satz 2. Man beachte, daß (f, g) wegen seiner Bilinearität (1., 1'. in **B**) und Satz 1 stetig ist.

[28]) Hierdurch wird die Möglichkeit, daß $\mathfrak{H}$ ein endlichvieldimensionaler Euklidischer Raum sei, ausgeschlossen. Die Bedingungen **A**, **B** sind übrigens in Anlehnung an ein (für endlich viele Dimensionen aufgestelltes) Axiomensystem des Vektorraumes von H. Weyl (Raum, Zeit, Materie, 5. Aufl., Berlin 1923, §§ 2—4) gebildet.

E. $\mathfrak{H}$ *ist vollständig. D. h.: wenn eine Folge* $f_1, f_2, \ldots$ *in* $\mathfrak{H}$ *der Cauchyschen Konvergenzbedingung genügt* (*zu jedem* $\varepsilon > 0$ *gibt es ein* $N = N(\varepsilon)$, *so daß aus* $m, n \geqq N$ $|f_m - f_n| \leqq \varepsilon$ *folgt*), *so ist sie konvergent* (*es existiert ein* f *aus* $\mathfrak{H}$, *so daß es zu jedem* $\varepsilon > 0$ *ein* $N = N(\varepsilon)$ *gibt, so daß aus* $n \geqq N$ $|f_n - f| \leqq \varepsilon$ *folgt*).

Im Anhang I werden wir zeigen, daß alle in der Einleitung I. und II. genannten Funktionenräume die Eigenschaften **A** bis **E** haben, also Spezialfälle des abstrakten Hilbertschen Raumes $\mathfrak{H}$ sind. Hier aber gehen wir daran zu beweisen, daß $\mathfrak{H}$ in der Tat durch die Eigenschaften **A** bis **E** vollständig charakterisiert ist: daß $\mathfrak{H}$ mit dem gewöhnlichen Hilbertschen Raume isomorph ist. Dazu ist es notwendig, einige allgemein bekannte Sätze über norm. orth. und vollst. norm. orth. Systeme zu beweisen.

Satz 3. *Jede norm. orth. Menge* $\mathfrak{M}$ *ist höchstens abzählbar, jede vollst. norm. orth. Menge* $\mathfrak{M}$ *ist genau abzählbar.*

Bemerkung. Wir schreiben darum von nun an alle norm. Orth. und vollst. norm. orth. Mengen als Folgen $\varphi_1, \varphi_2, \ldots$ (die im ersteren Falle eventuell abbrechen).

Beweis. $\mathfrak{M}$ sei norm. orth., für zwei beliebige f, g von $\mathfrak{M}$, $f \neq g$ gilt

$$|f - g|^2 = |f|^2 + |g|^2 - 2\,\Re(f, g) = 2, \qquad |f - g| = \sqrt{2};$$

wegen der Separabilität von $\mathfrak{H}$ (Bedingung **C**) ist also $\mathfrak{M}$ höchstens abzählbar.

Wir müssen noch zeigen, daß ein endliches norm. orth. System $\mathfrak{M}$ nicht vollst. ist. Seien die Elemente von $\mathfrak{M}$ $\varphi_1, \ldots, \varphi_n$. Unter den $a_1\varphi_1 + \ldots + a_n\varphi_n$ gibt es keine $n+1$ lin. unabh., also muß $\mathfrak{H}$ ein Element f haben, das von ihnen allen verschieden ist (Bedingung **D**). Dann ist

$$f^* = f - a_1\varphi_1 - \ldots - a_n\varphi_n$$

jedenfalls $\neq 0$, und zu allen $\varphi_1, \ldots, \varphi_n$ orth., wenn wir $a_k = (f, \varphi_k)$ $(k = 1, \ldots, n)$ setzen. Nach der Schlußbemerkung von Def. 3 ist also $\mathfrak{M}$ unvollst.

Satz 4. $\varphi_1, \varphi_2, \ldots$ *sei ein norm. orth. System. Dann ist die Reihe*

$$\sum_{n=1}^{\infty} (f, \varphi_n) \cdot \overline{(g, \varphi_n)}$$

für alle f, g *von* $\mathfrak{H}$ *absolut konvergent, insbesondere ist für* $f = g$

$$\sum_{n=1}^{\infty} |(f, \varphi_n)|^2 \leqq |f|^2.$$

Beweis. Für $a_n = (f, \varphi_n)$ gilt bekanntlich:

$$|f - \sum_{n=1}^{N} a_n \varphi_n|^2 = |f|^2 - 2\,\Re \sum_{n=1}^{N} (f, a_n \varphi_n) + \sum_{m,n=1}^{N} (a_m \varphi_m, a_n \varphi_n)$$
$$= |f|^2 - 2\,\Re \sum_{n=1}^{N} \overline{a_n} (f, \varphi_n) + \sum_{m,n=1}^{N} a_m \overline{a_n} (\varphi_m, \varphi_n)$$
$$= |f|^2 - 2 \sum_{n=1}^{N} |a_n|^2 + \sum_{n=1}^{N} |a_n|^2 = |f|^2 - \sum_{n=1}^{N} |a_n|^2 .$$

Da die linke Seite stets $\geqq 0$ ist, haben wir:

$$\sum_{n=1}^{N} |a_n|^2 \leqq |f|^2, \qquad \sum_{n=1}^{N} |(f, \varphi_n)|^2 \leqq |f|^2 .$$

Somit ist $\sum_{n=1}^{\infty} |(f, \varphi_n)|^2$ konvergent (absolut!), und zwar $\leqq |f|^2$. Und weil

$$|(f, \varphi_n) \cdot \overline{(g, \varphi_n)}| \leqq \frac{1}{2} |(f, \varphi_n)|^2 + \frac{1}{2} |(g, \varphi_n)|^2$$

ist, konvergiert auch $\sum_{n=1}^{\infty} (f, \varphi_n) \cdot \overline{(g, \varphi_n)}$ absolut. Damit ist alles bewiesen.

Satz 5. *$\varphi_1, \varphi_2, \ldots$ sei ein norm. orth. System. Die Reihe $\sum_{n=1}^{\infty} a_n \varphi_n$ konvergiert dann und nur dann, wenn $\sum_{n=1}^{\infty} |a_n|^2$ konvergiert.*

Beweis. Nach E genügt es zu zeigen, daß die Cauchysche Konvergenzbedingung für beide Reihen gleich lautet: Dies ist aber wegen $(m \geqq n)$

$$|\sum_{p=1}^{m} a_p \varphi_p - \sum_{p=1}^{n} a_p \varphi_p|^2 = |\sum_{p=n+1}^{m} a_p \varphi_p|^2 = \sum_{p,q=n+1}^{m} (a_p \varphi_p, a_q \varphi_q)$$
$$= \sum_{p,q=n+1}^{m} a_p \overline{a_q} (\varphi_p, \varphi_q) = \sum_{p=n+1}^{m} |a_p|^2 = \sum_{p=1}^{m} |a_p|^2 - \sum_{p=1}^{n} |a_p|^2$$

in der Tat der Fall.

Zusatz. *Für $f = \sum_{n=1}^{\infty} a_n \varphi_n$ gilt (im Falle der Konvergenz)*

$$(f, \varphi_n) = a_n .$$

Beweis. Für $N \geqq n$ ist

$$(\sum_{p=1}^{N} a_p \varphi_p, \varphi_n) = \sum_{p=1}^{N} a_p (\varphi_p, \varphi_n) = a_n ,$$

und wegen der Stetigkeit des inneren Produktes dürfen wir $N \to \infty$ lassen, was $(f, \varphi_n) = a_m$ ergibt.

Satz 6. *$\varphi_1, \varphi_2, \ldots$ sei ein norm. orth. System. Für jedes f ist die Reihe*

$$f' = \sum_{n=1}^{\infty} a_n \varphi_n , \qquad a_n = (f, \varphi_n) \qquad (n = 1, 2, \ldots)$$

konvergent, und zwar ist $f - f'$ zu allen $\varphi_1, \varphi_2, \ldots$ orth.

Beweis. Die Reihe konvergiert nach Satz 4 und 5, und nach dem Zusatz zu Satz 5 ist $(f', \varphi_n) = a_n = (f, \varphi_n)$, $(f - f', \varphi_n) = 0$; also ist alles bewiesen.

Satz 7. *Dazu, daß ein norm. orth. System $\varphi_1, \varphi_2, \ldots$ vollst. sei, ist eine jede der drei folgenden Bedingungen notwendig und hinreichend:*

α) *$\varphi_1, \varphi_2, \ldots$ spannen die abg. lin. M. $\mathfrak{H}$ auf.*

β) *Es ist für alle f*

$$f = \sum_{n=1}^{\infty} a_n \varphi_n, \quad a_n = (f, \varphi_n) \qquad (n = 1, 2, \ldots).$$

γ) *Es ist für alle f, g*

$$(f, g) = \sum_{n=1}^{\infty} (f, \varphi_n) \cdot \overline{(g, \varphi_n)}.$$

Beweis. Erstens folgt aus der Vollst. β): denn $f - f'$ aus Satz 6 muß verschwinden. Zweitens folgt aus β) α), denn f ist Häufungspunkt der $\sum_{n=1}^{N} a_n \varphi_n$, d. h. der von $\varphi_1, \varphi_2, \ldots$ aufgespannten lin. M. — also Element der abg. lin. M. Drittens folgt aus β) γ), denn es ist

$$\sum_{n=1}^{N} a_n \varphi_n \to f, \quad \left|\sum_{n=1}^{N} a_n \varphi_n\right|^2 \to |f|^2 \quad (\text{für } N \to \infty),$$

$$\left|\sum_{n=1}^{N} a_n \varphi_n\right|^2 = \sum_{m, n=1}^{N} a_m \overline{a_n} (\varphi_m, \varphi_n) = \sum_{n=1}^{N} |a_n|^2,$$

also

$$\sum_{n=1}^{\infty} |a_n|^2 = \sum_{n=1}^{\infty} |(f, \varphi_n)|^2 = |f|^2.$$

Wenn wir hier f durch $\frac{f+g}{2}$ und $\frac{f-g}{2}$ ersetzen und subtrahieren, gewinnen wir den Realteil von γ), wenn wir f, g durch $i f, g$ ersetzen, ihren Imaginärteil.

Viertens folgt aus α) die Vollst., denn sonst gäbe es ein zu allen $\varphi_1, \varphi_2, \ldots$ orth. $f \neq 0$. Also ist die ganze von $\varphi_1, \varphi_2, \ldots$ aufgespannte abg. lin. M., d. h. $\mathfrak{H}$, zu f orth. — also auch f: $|f|^2 = (f, f) = 0$, $f = 0$, entgegen der Annahme. Fünftens folgt auch aus γ) die Vollst.: denn für dasselbe f muß nach γ) $(f = g)$

$$|f|^2 = (f, f) = \sum_{n=1}^{\infty} 0^2 = 0, f = 0$$

sein.

Wir haben also das folgende logische Schema:

$$\begin{tikzcd}[row sep=tiny] & & \alpha \arrow[dr] & \\ \text{Vollständigkeit} \arrow[r] & \beta \arrow[ur] \arrow[dr] & & \text{Vollständigkeit.} \\ & & \gamma \arrow[ur] & \end{tikzcd}$$

Folglich sind diese vier Aussagen in der Tat gleichbedeutend.

Satz 8. *Zu jeder Folge $f_1, f_2, \ldots$ beliebiger Elemente von $\mathfrak{H}$ gibt es ein norm. orth. System $\varphi_1, \varphi_2, \ldots$, welches dieselbe lin. M. (also auch dieselbe abg. lin. M.) aufspannt (eine jede der beiden Folgen kann bereits im Endlichen abbrechen).*

Bemerkung. Wir wählen $f_1, f_2, \ldots$ überall dicht (Bedingung C!), dann spannt es natürlich die abg. lin. M. $\mathfrak{H}$ auf. Ebenso das nach Satz 8 konstruierte norm. orth. System. $\varphi_1, \varphi_2, \ldots$, nach Satz 7 α) ist es somit vollst. Damit ist die Existenz von vollst. norm. orth. Systemen erwiesen.

Beweis. Erstens können wir $f_1, f_2, \ldots$ durch eine Teilfolge von lauter lin. unabh. Elementen ersetzen, die dieselbe lin. M. aufspannt. In der Tat sei g_1 das erste $f_n \neq 0$, g_2 das erste $f_n \neq a_1 g_1$ (a_1 beliebig), g_3 das erste $f_n \neq a_1 g_1 + a_2 g_2$ (a_1, a_2 beliebig) usw. (wenn es kein $f_n \neq a_1 g_1 + \ldots + a_p g_p$ gibt, brechen wir mit g_p ab). $g_1, g_2, \ldots$ leistet offenbar das zunächst Geforderte.

Auf $g_1, g_2, \ldots$ wenden wir nunmehr das E. Schmidtsche „Orthogonalisationsverfahren“ an:

$$\begin{array}{ll} \gamma_1 = g_1, & \varphi_1 = \frac{1}{|\gamma_1|}\cdot\gamma_1, \\ \gamma_2 = g_2 - (g_2, \varphi_1)\cdot\varphi_1, & \varphi_2 = \frac{1}{|\gamma_2|}\cdot\gamma_2, \\ \gamma_3 = g_3 - (g_3, \varphi_1)\cdot\varphi_1 - (g_3, \varphi_2)\cdot\varphi_2, & \varphi_3 = \frac{1}{|\gamma_3|}\cdot\gamma_3, \\ \cdots\cdots\cdots & \cdots\cdots \end{array}$$

$\varphi_1, \varphi_2, \ldots$ spannen offenbar dieselbe lin. M. auf wie $g_1, g_2, \ldots$, d. h wie $f_1, f_2, \ldots$; ferner sieht man sofort, daß sie ein norm. orth. System bilden.

Auf Grund dieser Tatsachen können wir nun zeigen, daß $\mathfrak{H}$ eineindeutig auf den gewöhnlichen Hilbertschen Raum, d. i. die Menge aller Folgen komplexer Zahlen $(x_1, x_2, \ldots)$ mit endlichem $\sum_{n=1}^{\infty} |x_n|^2$, abgebildet werden kann, und zwar derart, daß ($\longleftrightarrow$ bezeichnet die Zuordnung) $f \longleftrightarrow (x_1, x_2, \ldots)$ die Folge $af \longleftrightarrow (a x_1, a x_2, \ldots)$ hat, $f \longleftrightarrow (x_1, x_2, \ldots)$, $g \longleftrightarrow (y_1, y_2, \ldots)$ die Folge $f + g \longleftrightarrow (x_1 + y_1, x_2 + y_2, \ldots)$, und außerdem $(f, g) = \sum_{n=1}^{\infty} x_n \overline{y_n}$.

Sei nämlich $\varphi_1, \varphi_2, \ldots$ ein (beliebiges, aber festes) vollst. norm. orth. System, dann ordnen wir f die Folge $(x_1, x_2, \ldots)$ mit $x_n = (f, \varphi_n)$ ($n = 1, 2, \ldots$) zu. Daß $(x_1, x_2, \ldots)$ zum Hilbertschen Raume gehört, folgt aus Satz 4, daß zu jedem solchen $(x_1, x_2, \ldots)$ ein f existiert, aus Satz 5 und dessen Zusatz. Die Ein-eindeutigkeit folgt aus der Vollst. (wenn stets $(f, \varphi_n) = (g, \varphi_n)$ ist, so ist $f - g$ zu allen φ_n orth., also $= 0$). Von den drei Schlußbehauptungen sind die zwei ersten klar, die dritte folgt aus Satz 7 γ).

Somit ist gezeigt, daß $\mathfrak{H}$ dem gewöhnlichen Hilbertschen Raume isomorph ist, also zwei **A** bis **E** genügende Systeme $\mathfrak{H}$ einander; d. h. **A** bis **E** sind ein „kategorisches" Axiomensystem, sie legen alle Eigenschaften von $\mathfrak{H}$ fest.

II. Operatoren in $\mathfrak{H}$.

Ein Operator ist eine Funktion, die in einer Teilmenge von $\mathfrak{H}$ definiert ist (eventuell in ganz $\mathfrak{H}$) und Werte aus $\mathfrak{H}$ annimmt. Wir wollen zunächst einige besonders wichtige Sorten von Operatoren definieren.

Definition 5. Ein Operator R ist linear (kurz: lin.), wenn bei sinnvollen $Rf_1, \ldots, Rf_n$ auch $R(a_1 f_1 + \ldots + a_n f_n)$ sinnvoll, und zwar $= a_1 Rf_1 + \ldots + a_n Rf_n$ ist. Ein Operator R ist abgeschlossen (kurz: abg.), wenn für jede Folge $f_1, f_2, \ldots$, für die alle Rf_n sinnvoll sind, die Konvergenzen $f_n \to f$, $Rf_n \to f^*$ die Konsequenz haben, daß Rf sinnvoll und zwar $= f^*$ ist[29]).

Definition 6. Ein Operator R ist Hermitesch (kurz: ein H.O.), wenn erstens sein Definitionsbereich die abg. lin. M. $\mathfrak{H}$ aufspannt[30]), und zweitens (soweit Rf, Rg Sinn haben)

$$(f, Rg) = (Rf, g)$$

gilt. Ein Operator R ist längentreu, wenn er abg. lin. ist und (soweit Rf Sinn hat) $|f| = |Rf|$ gilt.

Wir beginnen damit, daß wir zwei Sätze über H. O. und längentreue O. beweisen. Wir erinnern noch daran, daß wir einen Operator S Fortsetzung eines Operators R nennen (kurz: Forts.), wenn überall wo Rf Sinn hat, auch Sf sinnvoll und zwar $= Rf$ ist — und eigentliche (kurz: eig.) Forts., wenn $R \neq S$ ist, d. h. S an mehr Stellen Sinn hat, wie R.

Satz 9. *Sei R ein H. O. Es gibt einen lin. H. O. $\hat{R}$, der Forts. von R ist und von dem alle ebensolchen H. O. ihrerseits Forts. sind; und es gibt einen abg. lin. H. O. $\tilde{R}$, der Forts. von R (also auch von $\hat{R}$) ist, und von dem alle ebensolchen H. O. ihrerseits Forts. sind.*

Beweis. Wenn wir $\hat{R}$ so definieren könnten: $\hat{R}f$ hat Sinn, wenn $f = a_1 f_1 + \ldots + a_n f_n$ ist, $Rf_1, \ldots, Rf_n$ sinnvoll, und zwar ist es dann $= a_1 Rf_1 + \ldots + a_n Rf_n$; und $\tilde{R}$ so: $\tilde{R}f$ hat Sinn, wenn eine Folge

---

[29]) Dies ist offenbar eine stetigkeits-ähnliche Eigenschaft, und zwar in jedem „kompakten" topologischen Raume der Stetigkeit gleichwertig. Im Hilbertschen Raume bedeutet sie aber viel weniger, für H. O. ist sie im wesentlichen immer da, vgl. Satz 9.

[30]) So weit schwächen wir (mit Rücksicht auf die Anwendung auf Matrizen, vgl. Anhang III) die Bedingung γ) von Einleitung V ab: dort wurde Überalldichtheit des Definitionsbereiches verlangt.

$g_1, g_2, \ldots$ existiert mit lauter sinnvollen $\hat{R}g_n$, wobei $g_n \to f$, $\hat{R}g_n \to f^*$, und zwar ist es dann $= f^*$ — dann würden diese $\hat{R}$, $\tilde{R}$ offenbar alles Gewünschte leisten. Es fragt sich aber, ob diese Festsetzungen widerspruchfrei möglich sind.

Sie sind es gewiß, wenn $a_1 f_1 + \ldots + a_m f_m = b_1 g_1 + \ldots + b_n g_n$ ($Rf_1, \ldots, Rf_m$, $Rg_1, \ldots, Rg_n$ sinnvoll) $a_1 Rf_1 + \ldots + a_m Rf_m = b_1 Rg_1 + \ldots + b_n Rg_n$ nach sich zieht; und ebenso $f_n \to f$, $g_n \to f$ (alle $\hat{R}f_n$, $\hat{R}g_n$ sinnvoll), $\hat{R}f_n \to f^*$, $\hat{R}g \to g^*$ die Konsequenz $f^* = g^*$ hat. Es genügt offenbar zu beweisen: Aus $a_1 f_1 + \ldots + a_n f_n = 0$ ($Rf_1, \ldots, Rf_n$ sinnvoll), folgt $a_1 Rf_1 + \ldots + a_n Rf_n = 0$, aus $f_n \to 0$ (alle $\hat{R}f_n$ sinnvoll), $\hat{R}f_n \to f^*$ folgt $f^* = 0$.

Gelte zunächst die erste Prämisse. Wenn Rg sinnvoll ist, so ist

$$(g, a_1 Rf_1 + \ldots + a_n Rf_n) = \sum_{p=1}^{n} (g, a_p Rf_p) = \sum_{p=1}^{n} (Rg, a_p f_p)$$
$$= (Rg, a_1 f_1 + \ldots + a_n f_n) = 0.$$

Alle diese g stehen also zu $a_1 Rf_1 + \ldots + a_n Rf_n$ orth., also auch ihre abg. lin. M., $\mathfrak{H}$ — also ist $a_1 Rf_1 + \ldots + a_n Rf_n$ auch zu sich selbst orth., d. h. $= 0$. Betrachten wir nunmehr die zweite Prämisse. Wenn $\hat{R}g$ Sinn hat, so ist ($\hat{R}$ ist ja ein H. O.)

$$(g, f^*) = (g, \text{limes}\, \hat{R}f_n) = \text{limes}\, (g, \hat{R}f_n) = \text{limes}\, (\hat{R}g, f_n)$$
$$= (\hat{R}g, \text{limes}\, f_n) = 0.$$

Somit stehen alle diese g zu f^* orth., also auch ihre abg. lin. M., $\mathfrak{H}$ — und dies hat wieder $f^* = 0$ zur Folge.

Satz 10. *Sei U ein längentreuer O. Dann ist sowohl der Definitionsbereich als auch der Wertevorrat von U je eine abg. lin. M., und U bildet sie stetig und ein-eindeutig aufeinander ab. Soweit U Sinn hat, gilt* $(f, g) = (Uf, Ug)$.

Bemerkung. Wenn Definitionsbereich wie Wertevorrat beide $= \mathfrak{H}$ sind, so heiße U, im Sinne der üblichen Terminologie, unitär.

Beweis. Da U lin. ist, sind Definitionsbereich und Wertevorrat lin. M. Weiter ist

$$|Uf - Ug| = |U(f-g)| = |f-g|,$$

also hat $Uf = Ug$ die Folge $f = g$, d. h. die Abbildung ist ein-eindeutig, ferner folgt hieraus ihre und ihrer Inversen Stetigkeit. Dies verursacht, zusammen mit der Abg. von U, daß Definitionsbereich und Wertevorrat beide abg. Mengen sind.

Wenn wir in $|f|^2 = |Uf|^2$ f durch $\frac{f+g}{2}$ und $\frac{f-g}{2}$ ersetzen und subtrahieren, so wird $\Re(f, g) = \Re(Uf, Ug)$. Wenn wir f, g durch if, g er-

setzen, erhalten wir dasselbe für die Imaginärteile. Also ist $(f, g) = (Uf, Ug)$. Damit ist alles bewiesen. —

Einige Eigenschaften von $\tilde{R}$ (R ein H. O.) sind evident: $\tilde{\tilde{R}} = \tilde{R}$, wenn S Forts. von R ist, so ist $\tilde{S}$ Forts. von $\tilde{R}$. Wir definieren nun die Maximalität von H. O.

Definition 7. Ein H. O. R heiße maximal (kurz: max.), wenn es keine eig. Forts. von R gibt (die auch H. O. ist!).

Da $\tilde{R}$ Forts. von R ist, muß für max. R $R = \tilde{R}$ sein. Ferner sei R ein H. O., S max. H. O., S Forts. von R; dann ist $S = \tilde{S}$ auch Forts. von $\tilde{R}$. D. h. jeder max. H. O. ist abg. lin., und ein H. O. R hat genau dieselben max. Forts., wie der abg. lin. H. O. $\tilde{R}$. Wir werden in den späteren Kapiteln die Forts. beliebiger H. O. zu max. untersuchen (vgl. auch die Einleitung VI., VII.), das soeben Gesagte zeigt, daß wir uns dabei durchweg auf abg. lin. Operatoren beschränken dürfen.

Wir definieren weiter:

Definition 8. Sei R ein H. O. f ist ein Erweiterungselement von R, und f^* wird f durch R zugeordnet, wenn für alle sinnvollen Rg $(f, Rg) = (f^*, g)$ gilt. f, f^* gehören zur $+$, 0, $-$-Klasse, je nachdem ob $\mathfrak{J}(f^*, f) >, =, < 0$ ist.

Satz 11. *f^* ist durch f eindeutig bestimmt (vgl. Def. 8); wenn Rf Sinn hat, so ist f Erweiterungselement von der 0-Klasse, und zwar $f^* = Rf$. R ist dann und nur dann max., wenn es keine weiteren Erweiterungselemente von der 0-Klasse gibt.*

Beweis. Wären f^*, f^{**} beide f zugeordnet, so müßte $(f^*, g) = (f^{**}, g)$, $(f^* - f^{**}, g) = 0$ sein für alle sinnvollen Rg, also auch für die ganze abg. lin. M. dieser g, $\mathfrak{H}$. Hieraus folgern wir wieder $f^* - f^{**} = 0$. Die zweite Behauptung ist klar. Bei der dritten ist die Hinreichendheit der Bedingung klar, die Notwendigkeit folgt daraus, daß jedes Erweiterungselement der 0-Klasse zur eig. Forts. von R benutzt werden könnte.

Die durch Satz 11 gegebene Charakterisierung der Max. legt nun die folgende Begriffsbildung nahe.

Definition 9. Ein H. O. R heiße hypermaximal (kurz: hypermax.), wenn es keine anderen Erweiterungselemente von R gibt als die, für welche R Sinn hat. (D. h. wenn R max. ist und die $+$- und $-$-Klassen leer sind.)

Die Wichtigkeit dieses Begriffes[31]) wird sich noch zeigen, vgl. z. B. Satz 36.

[31]) Er stammt von Herrn E. Schmidt, vgl. Anm. [23]). Bei Beschränkung auf den reellen Hilbertschen Raum tritt die Unterscheidung max. — hypermax. nicht auf.

Wir beweisen noch einen Satz über die Stetigkeit von lin. O.

Satz 12. *Für einen lin. O. R ist eine jede der folgenden Aussagen mit der Stetigkeit gleichwertig.*

$$\alpha)\quad |Rf| \leqq C\cdot|f|, \qquad \beta)\quad |(f, Rg)| \leqq C\cdot|f|\ |g|.$$

Wenn R ein H. O. ist, so kommt noch hinzu:

$$\gamma)\quad |(f, Rf)| \leqq C\cdot|f|^2$$ [32]).

Bemerkung. Nach Hilbert sagen wir bei lin. H. O. für stetig auch beschränkt. Bei nicht-lin. O. unterlassen wir es, die Stetigkeit zu untersuchen. (Die Hilbertsche Definition ist β).)

Beweis. Wegen der Lin. folgt aus α) $|Rf - Rg| = |R(f-g)| \leqq C|f-g|$, d. h. die Stetigkeit. In β) brauchen wir nur $f = Rg$ zu setzen, um $|Rg|^2 \leqq C\cdot|Rg|\cdot|g|$, $|Rg| \leqq C\cdot|g|$, d. h. α), zu gewinnen. Wenn R stetig ist, so gibt es wegen $R0 = 0$ (Lin.!) ein $\varepsilon > 0$, so daß $|f| \leqq \varepsilon$ $|Rf| \leqq 1$ nach sich zieht. Daher ist stets (Lin.!) $|Rf| \leqq \frac{1}{\varepsilon}\cdot|f|$, also $|(f, Rg)| \leqq |f|\cdot|Rg| \leqq \frac{1}{\varepsilon}\cdot|f|\cdot|g|$, d. h. β) gilt. Somit sind α), β) der Stetigkeit wirklich gleichwertig.

γ) besagt offenbar weniger als β), also genügt es β) für H. O. aus γ) zu folgern. Das geschieht ähnlich wie bei Satz 1. Wir setzen für f $\frac{f+g}{2}$ und $\frac{f-g}{2}$ ein und subtrahieren: dann wird

$$\Re(f, Rg) \leqq C\cdot\left(\frac{1}{2}|f|^2 + \frac{1}{2}|g|^2\right).$$

Indem wir (wie bei Satz 1) f, g durch $af, \frac{1}{a}g$ $(a > 0)$ ersetzen und das Minimum der rechten Seite nehmen, sodann f, g durch $\vartheta f, g$ $(|\vartheta| = 1)$, und das Maximum der linken bilden, erhalten wir $|(f, Rg)| \leqq C\cdot|f|\cdot|g|$, d. h. β). Aber hier ist Rf, Rg als sinnvoll vorausgesetzt. Von der Sinnvollheit von Rf befreien wir uns: in β) kommt Rf nicht vor, es gilt für alle f, für die dies Sinn hat, also in einer überall dichten Menge (weil R ein lin. H. O. ist); da beide Seiten stetig sind, gilt es also für alle f überhaupt.

Für lin. H. O. haben wir als Bedingung der Beschränktheit:

$$-C\cdot|f|^2 \leqq (f, Rf) \leqq C\cdot|f|^2$$

(Satz 12 γ)). Dies zerlegen wir wie üblich:

---

[32]) Für einen H. O. muß jedes (f, Rf) reell sein, da ja $(f, Rf) = (Rf, f) = \overline{(f, Rf)}$ gilt.

Definition 10. Ein lin. H. O. R heiße nach oben bzw. unten halbbeschränkt, wenn stets

$$(f, Rf) \leqq C \cdot |f|^2 \quad \text{bzw.} \quad \geqq -C \cdot |f|^2$$

gilt. (Beides zusammen ist nach Satz 12 γ) Beschränktheit.) Wenn insbesondere $(f, Rf) \geqq 0$ gilt, so heiße er definit.

III. Lineare Mannigfaltigkeiten und Projektionsoperatoren.

Definition 11. $\mathfrak{M}, \mathfrak{N}$ seien zwei, nicht notwendig abg., lin. M. Dann sind $\mathfrak{M} + \mathfrak{N}$ und $\mathfrak{M} \times \mathfrak{N}$, die von ihrer Vereinigungsmenge aufgespannte lin. M. und ihr Durchschnitt, wieder lin. M. Wenn insbesondere $\mathfrak{M}, \mathfrak{N}$ abg. lin. M. sind, so ist es $\mathfrak{M} \times \mathfrak{N}$ offenbar ebenfalls. (Ob $\mathfrak{M} + \mathfrak{N}$ auch abg. ist, ist ungewiß; für zueinander orth. $\mathfrak{M}, \mathfrak{N}$ folgt es aus Satz 17.) Im folgenden sollen $\mathfrak{M}, \mathfrak{N}$ als abg. lin. M. vorausgesetzt werden. Wenn $\mathfrak{M}$ Teilmenge von $\mathfrak{N}$ ist, so ist $\mathfrak{N} - \mathfrak{M}$ ihre Differenz, d. h. die Menge aller Elemente von $\mathfrak{N}$, die zu allen Elementen von $\mathfrak{M}$ orth. sind, auch eine abg. lin. M. $\mathfrak{H} - \mathfrak{M}$ heiße auch die Komplementäre von $\mathfrak{M}$.

Satz 13. *Sei $\mathfrak{M}$ eine abg. lin. M. Jedes f kann auf eine und nur eine Art in $f = g + h$, g von $\mathfrak{M}$, h von $\mathfrak{H} - \mathfrak{M}$ zerlegt werden.*

Bemerkung. g heißt dann die Projektion von f in $\mathfrak{M}$, die Zuordnung von g zu f ist offenbar ein überall sinnvoller Operator, wir nennen ihn den Projektionsoperator von $\mathfrak{M}$, $P_{\mathfrak{M}}$ ($\mathfrak{M}$ ist eine abg. lin. M.!).

Beweis. Die Eindeutigkeit ist klar: Aus $g_1 + h_1 = g_2 + h_2$ (g_1, g_2 von $\mathfrak{M}$, h_1, h_2 von $\mathfrak{H} - \mathfrak{M}$) folgt $g_1 - g_2 = h_2 - h_1$, dieses gehört also zu $\mathfrak{M}$ wie zu $\mathfrak{H} - \mathfrak{M}$, ist somit zu sich selbst orth., also $= 0$; d. h. $g_1 = g_2$, $h_1 = h_2$. Es bleibt übrig, die Existenz zu beweisen.

Da $\mathfrak{H}$ separabel ist, gibt es eine in $\mathfrak{M}$ überall dichte Folge $f_1, f_2, \ldots$ [33]), diese spannt also die abg. lin. M. $\mathfrak{M}$ auf. Nach Satz 8 gibt es also ein norm. orth. System $\varphi_1, \varphi_2, \ldots$, welches dasselbe tut. Nach Satz 6 konvergiert $g = \sum_{n=1}^{\infty} (f, \varphi_n) \cdot \varphi_n$, und es gehört mit den $\varphi_1, \varphi_2, \ldots$ zu $\mathfrak{M}$, und $h = f - g$ ist zu allen $\varphi_1, \varphi_2, \ldots$ orth. Also auch zu ihrer abg. lin. M., $\mathfrak{M}$ — d. h. h gehört zu $\mathfrak{H} - \mathfrak{M}$. Damit ist alles bewiesen.

Satz 14. *Der Projektionsoperator $P_{\mathfrak{M}}$ ist ein überall sinnvoller max. H. O., es ist $P_{\mathfrak{M}}^2 = P_{\mathfrak{M}}$.*

Beweis. Daß $P_{\mathfrak{M}}$ überall sinnvoll ist, wissen wir. Er ist ein H. O., wenn wir $(f, P_{\mathfrak{M}} g) = (P_{\mathfrak{M}} f, g)$ beweisen können, d. h. für $f = h + k$,

---

[33]) Jede Teilmenge einer separablen topologischen Menge ist wieder separabel (siehe z. B. Hausdorff, Einführung in die Mengenlehre, Leipzig-Berlin 1914).

$g = h' + k'$ (h, h' von $\mathfrak{M}$, k, k' von $\mathfrak{H} - \mathfrak{M}$) $(f, h') = (h, g)$. Nun ist $(k, h') = 0$, $(h, k') = 0$, also

$$(f, h') = (h + k, h') = (h, h') = (h, h' + k') = (h, g).$$

Wenn f zu $\mathfrak{M}$ gehört, ist die Zerlegung von Satz 13 $f = f + 0$, also $P_{\mathfrak{M}} f = f$. Da $P_{\mathfrak{M}} f$ zu $\mathfrak{M}$ gehört, ist also $P_{\mathfrak{M}}(P_{\mathfrak{M}} f) = P_{\mathfrak{M}} f$, d. h. $P_{\mathfrak{M}}^2 = P_{\mathfrak{M}}$. Da $P_{\mathfrak{M}}$ überall sinnvoll ist, muß es max. sein. —

Die max. H. O. E mit der Eigenschaft $E^2 = E$ haben aber für uns ein selbständiges Interesse (sie entsprechen den „Einzelformen" bei Hilbert), wir untersuchen sie darum zunächst für sich.

Satz 15. *Zu jedem max. H. O. E mit der Eigenschaft $E^2 = E$* [34]) *gibt es eine abg. lin. M. $\mathfrak{M}$, so daß $P_{\mathfrak{M}} = E$ ist. $\mathfrak{M}$ ist der Wertevorrat von E.*

Beweis. Da $\mathfrak{M}$ offenbar der Wertevorrat von $P_{\mathfrak{M}}$ ist, genügt es zu zeigen, daß $P_{\mathfrak{M}} = E$ ist, wenn $\mathfrak{M}$ der Wertevorrat von E ist.

Wegen $E^2 = E$ ist, wie man leicht einsieht, $\mathfrak{M}$ die Menge aller f mit $Ef = f$, also eine abg. lin. M. (da E max. ist, ist es auch lin. abg.!). Wenn Ef Sinn hat, so ist $f - Ef$ zu jedem Element von $\mathfrak{M}$, d. h. jedem Eg, orth.:

$$(Eg, f - Ef) = (g, Ef - E^2 f) = 0;$$

also gehört es zu $\mathfrak{H} - \mathfrak{M}$. Daher ist $P_{\mathfrak{M}} f = Ef$, also $P_{\mathfrak{M}}$ Forts. von E. Da E max. ist, ist folglich $P_{\mathfrak{M}} = E$.

Definition 12. Einen Operator, der den Bedingungen von Satz 15 genügt, nennen wir Projektionsoperator (kurz: P. O., nach Satz 15 sind die P. O. mit den $P_{\mathfrak{M}}$ identisch und den abg. lin. M. $\mathfrak{M}$ ein-eindeutig zugeordnet). Eine abg. lin. M. $\mathfrak{M}$ und ihren P. O. $P_{\mathfrak{M}}$ nennen wir zueinandergehörig. Man sieht sofort: zu den abg. lin. M. (0) [35]) und $\mathfrak{H}$ gehören die P. O. 0 bzw. 1 [36]), ferner ist $P_{\mathfrak{H} - \mathfrak{M}} = 1 - P_{\mathfrak{M}}$. (Also sind $0, 1$ P. O., und mit E ist es auch $1 - E$.)

Satz 16. *Ein P. O. E ist stets beschränkt und definit, es ist nämlich*

$$(f, Ef) = |Ef|^2, \quad |Ef| \leqq |f|.$$

Beweis. Es ist $(f, Ef) = (f, E^2 f) = (Ef, Ef) = |Ef|^2$, also ist $(f, Ef) \geqq 0$. Wenden wir dies auf den P. O. $1 - E$ an, so wird: $(f, (1 - E) f) \geqq 0$, $(f, Ef) \leqq (f, f) = |f|^2$, also $|Ef|^2 \leqq |f|^2$, $|Ef| \leqq |f|$. Damit sind die Formeln der zweiten Zeile bewiesen. Die erste Zeile folgt aus ihnen.

---

[34]) D. h. wenn Ef sinnvoll ist, so ist es auch $E(Ef)$, und zwar $= Ef$.

[35]) (0) ist die Menge, deren einziges Element die 0 (aus $\mathfrak{H}$) ist. Die leere Menge sehen wir nicht als lin. M. an.

[36]) $0, 1$ sind zwei überall sinnvolle Operatoren, durch $0f = 0$, $1f = f$ definiert.

Satz 17. *$\mathfrak{M}$, $\mathfrak{N}$ seien zwei abg. lin. M., E, F die zugehörigen P. O. EF ist dann und nur dann ein P. O., wenn $EF = FE$ ist, d. h. E, F vertauschbar; dann nennen wir auch $\mathfrak{M}$, $\mathfrak{N}$ vertauschbar, EF ist der P. O. von $\mathfrak{M} \times \mathfrak{N}$. $E + F$ ist dann und nur dann ein P. O., wenn $EF = 0$ ist (oder auch: wenn $FE = 0$ ist); dies bedeutet, daß ganz $\mathfrak{M}$ auf ganz $\mathfrak{N}$ orth. steht, wir nennen dann auch E, F orth., $E + F$ ist der P. O. von $\mathfrak{M} + \mathfrak{N}$. $E - F$ ist dann und nur dann ein P. O., wenn $EF = F$ ist (oder auch: wenn $FE = F$ ist); dies bedeutet, daß $\mathfrak{N}$ Teilmenge von $\mathfrak{M}$ ist, wir nennen dann auch F Teil von E, $E - F$ ist der P. O. von $\mathfrak{M} - \mathfrak{N}$.*[37])

Beweis. EF ist überall sinnvoll, also ein max. H. O., wenn es der Symmetrie-Bedingung $(EFf, g) = (f, EFg)$ genügt. Nun ist

$$(f, EFg) = (Ef, Fg) = (FEf, g), \quad (EFf, g) - (f, EFg) = (\{EF - FE\} f, g),$$

damit dies für alle g verschwinde, muß $(EF - FE) f = 0$ sein (für alle f), also $EF - FE = 0$. Unsere Bedingung ist also bereits für den max. H. O.-Charakter von EF notwendig und hinreichend. Und da sie

$$(EF)^2 = EFEF = EEFF = E^2 F^2 = EF$$

zur Folge hat, auch für den P. O.-Charakter.

Für $E + F$ ist der H. O.-Charakter klar, ebenso die Max. (es ist überall sinnvoll!), es genügt also $(E + F)^2 = E + F$ zu untersuchen. Wegen

$$(E + F)^2 = E^2 + F^2 + EF + FE = (E + F) + (EF + FE)$$

bedeutet es $EF + FE = 0$. Hieraus folgt aber $EF = 0$, denn man multipliziere links mit E: $EF + EFE = 0$, und dieses rechts mit E: $2 \cdot EFE = 0$ beides zusammen ergibt $EF = 0$[38]). Ist umgekehrt $EF = 0$, so ist es ein P. O., also $FE = EF = 0$, $EF + FE = 0$. Somit ist $EF = 0$ wirklich charakteristisch, Vertauschen von E, F zeigt, daß es $FE = 0$ gleichfalls ist.

$E - F$ ist dann und nur dann ein P. O., wenn $1 - (E - F) = (1 - E) + F$ einer ist, da $1 - E$ und F solche sind, bedeutet dies: $(1 - E) F = 0$, $F - EF = 0$, oder auch: $F(1 - E) = 0$, $F - FE = 0$.

Es bleibt noch übrig, die Aussagen über $\mathfrak{M}$, $\mathfrak{N}$, $\mathfrak{M} \times \mathfrak{N}$, $\mathfrak{M} \pm \mathfrak{N}$ zu beweisen. Sei erstens $EF = FE$. Jedes $EFf = FEf$ gehört zu $\mathfrak{M}$ und $\mathfrak{N}$, also zu $\mathfrak{M} \times \mathfrak{N}$; wenn aber f zu $\mathfrak{M} \times \mathfrak{N}$, d. h. $\mathfrak{M}$ und $\mathfrak{N}$, gehört, so ist $EFf = Ef = f$ — d. h. $\mathfrak{M} \times \mathfrak{N}$ ist der Wertevorrat von EF, also dieses sein P. O. Sei zweitens $EF = 0$ (also $FE = 0$). Wenn f zu $\mathfrak{M}$, g zu $\mathfrak{N}$ gehört, so ist $(f, g) = (Ef, Fg) = (f, EFg) = 0$, d. h. $\mathfrak{M}$ zu $\mathfrak{N}$ orth., ebenso klar ist die Umkehrung. Für jedes f gehört Ef zu $\mathfrak{M}$, Ff zu $\mathfrak{N}$,

---

[37]) Da es sich um überall sinnvolle Operatoren handelt, dürfen wir ruhig EF, $E \pm F$ bilden; sonst wären die Definitionsbereiche dieser Operatoren noch genauer zu erörtern.

[38]) Dieser Kunstgriff stammt von Herrn E. Schmidt.

also $(E+F)f$ zu $\mathfrak{M}+\mathfrak{N}$; wenn f zu $\mathfrak{M}+\mathfrak{N}$ gehört, so ist es $=g+h$, g von $\mathfrak{M}$, h von $\mathfrak{N}$, also

$$(E+F)f=(E+F)(g+h)=Eg+Fg+Eh+Fh=g+0+0+h=f$$

— d. h. $\mathfrak{M}+\mathfrak{N}$ ist der Wertevorrat von $E+F$, also eine abg. lin. M., und dieses sein P. O. (Satz 15). Sei drittens $EF=F$ (also $FE=F$). Dies bedeutet $EFf=Ff$ für alle f, d. h. $Eg=g$ für alle g von $\mathfrak{N}$, d. h. (da dies die Definitionsgleichung von $\mathfrak{M}$ ist, vgl. den Beweis von Satz 15) daß $\mathfrak{N}$ Teilmenge von $\mathfrak{M}$ ist. Wegen $E-F=E-EF=E(1-F)$ ist $E-F$ der P. O. von $\mathfrak{M}\times(\mathfrak{H}-\mathfrak{N})=\mathfrak{M}-\mathfrak{N}$.

Damit sind alle unsere Behauptungen bewiesen.

Satz 18. *E ist dann und nur dann Teil von F, wenn stets*

$$|Ef|\leqq|Ff|$$

gilt. $E_1+\ldots+E_p$ ist dann und nur dann P. O. ($E_1,\ldots,E_p$ P. O.), wenn

$$E_mE_n=0 \qquad (m \neq n;\ m,n=1,\ldots,p)$$

gilt.

Beweis. Wenn E Teil von F ist, so ist $E=EF$, also $|Ef|=|EFf|\leqq|Ff|$. Gilt dies umgekehrt für alle f, so hat $f=Ef$ die Konsequenz

$$|Ff|\geqq|Ef|=|f|, \qquad (f,Ff)=|Ff|^2\geqq|f|^2=(f,f),$$
$$|(1-F)f|^2=(f,(1-F)f)\leqq 0, \qquad (1-F)f=0,$$

also $f=Ff$. D. h. wenn E, F die P. O. von $\mathfrak{M}, \mathfrak{N}$ sind, so ist $\mathfrak{M}$ Teil von $\mathfrak{N}$ — also E Teil von F.

Daß $E_mE_n=0$ $(m\neq n)$ für den P. O.-Charakter von $E_1+\ldots+E_p$ hinreicht, folgt durch wiederholte Anwendung von Satz 17; seine Notwendigkeit erkennen wir so. Es ist $(m\neq n)$

$$|E_mf|^2+|E_nf|^2\leqq|E_1f|^2+\ldots+|E_pf|^2=(f,E_1f)+\ldots+(f,E_pf)$$
$$=(f,(E_1+\ldots+E_p)f)\leqq|f|^2;$$

also für $E_nf=f$ $|E_mf|^2\leqq 0$, $E_mf=0$. Da für $f=E_ng$ das erstere zutrifft, ist stets $E_mE_ng=0$, d. h. $E_mE_n=0$, wie behauptet wurde. —

Zum Schluß noch ein Satz über die Konvergenz von P. O.

Satz 19. *$E_1, E_2, \ldots$ sei eine Folge von P. O., und entweder E_m Teil von E_n für alle $m<n$, oder E_n Teil von E_m für alle $m<n$. Dann gibt es einen P. O. E, so daß für jedes f $E_nf\to Ef$ gilt.*

Beweis. Wenn wir $E_1, E_2, \ldots, E$ durch $1-E_1, 1-E_2, \ldots, 1-E$ ersetzen, geht der erste Unterfall in den zweiten über[39]), es genügt daher etwa den ersteren zu betrachten.

[39]) Denn wenn E Teil von F ist, so ist $1-F$ Teil von $1-E$, z. B. darum, weil $(1-E)(1-F)=1-E-F+EF=1-F$ ist.

Mit wachsendem n wächst $|E_n f|^2$ (Satz 18), und es bleibt $\leqq |f|^2$ (Satz 16), daher existiert zu jedem $\varepsilon > 0$ ein $N = N(\varepsilon)$, so daß für jedes $m, n \geqq N$ $|E_n f|^2 - |E_m f|^2 \leqq \varepsilon$ bleibt. Sei noch $m < n$, dann ist $E_n - E_m$ ein P. O., also

$$|E_n f - E_m f|^2 = (f, (E_n - E_m) f) = (f, E_n f) - (f, E_m f) = |E_n f|^2 - |E_m f|^2 \leqq \varepsilon.$$

Somit hat die Folge $E_1 f, E_2 f, \ldots$ einen Limes, er heiße Ef.

E ist ein überall sinnvoller Operator. Da

$$(f, Eg) = \operatorname{limes}(f, E_n g) = \operatorname{limes}(E_n f, g) = (Ef, g)$$

ist, ist es ein max. H. O. Wir brauchen daher nur noch $E^2 = E$ zu beweisen. Es ist

$$(f, E^2 g) = (Ef, Eg) = \operatorname{limes}(E_n f, E_n g) = \operatorname{limes}(f, E_n^2 g) = \operatorname{limes}(f, E_n g)$$
$$= (f, Eg),$$

also in der Tat $E^2 = E$.

IV. Reduzibilität von Operatoren.

Bei endlichvieldimensionalen Matrizen versteht man unter Reduzibilität bekanntlich die Möglichkeit, die Matrix unitär auf eine solche Form zu transformieren, daß in ihr (die etwa $N = M_1 + M_2$-dimensional sein mag) nur die gemeinsamen Felder der M_1 ersten Zeilen und Spalten und die gemeinsamen Felder der M_2 letzten Zeilen und Spalten anders als mit Nullen besetzt sein können. Wenn die Matrix als lineare Transformation aufgefaßt wird, so heißt dies: der Raum ist in zwei (zueinander orth.) lin. M. von M_1- bzw. M_2-Dimensionen zerlegbar, deren jede durch die Transformation in ein Teil von sich übergeführt wird.

Bei den Operatoren in $\mathfrak{H}$ analogisieren wir dies folgendermaßen, wobei wir uns auf abg. lin. Operatoren beschränken:

Definition 13. Ein abg. lin. Operator R heiße durch die abg. lin. M. $\mathfrak{M}$ reduziert, wenn aus der Sinnvollheit von Rf auch die von $RP_{\mathfrak{M}}f$ folgt, und zwar dieses $= P_{\mathfrak{M}} Rf$ ist[40]).

Durch (0) und $\mathfrak{H}$ wird offenbar jedes R reduziert, und wenn $\mathfrak{M}$ das R reduziert, so tut es auch $\mathfrak{H} - \mathfrak{M}$ (wegen $P_{\mathfrak{H}-\mathfrak{M}} = 1 - P_{\mathfrak{M}}$). Irreduzibel heiße R, wenn es nur durch (0) und $\mathfrak{H}$ reduziert wird.

Wenn R durch $\mathfrak{M}$ reduziert wird, so gehört für jedes f von $\mathfrak{M}$ offenbar auch Rf zu $\mathfrak{M}$ ($Rf = RP_{\mathfrak{M}}f = P_{\mathfrak{M}}Rf$), also erzeugt R auch einen Operator in $\mathfrak{M}$ (d. h. eine Funktion, deren Definitionsbereich und Wertevorrat Teilmengen von $\mathfrak{M}$ sind). Diesen Operator nennen wir den in $\mathfrak{M}$ liegenden Teil von R.

---

[40]) Dies kann also auch als Vertauschbarkeit von R mit $P_{\mathfrak{M}}$ aufgefaßt werden.

Wir beweisen nun zwei Sätze über das Reduzieren von Operatoren.

Satz 20. *R sei ein abg. lin. Operator, $\mathfrak{M}$ eine abg. lin. M. Damit $\mathfrak{M}$ R reduziere, ist notwendig, daß mit Rf auch $RP_{\mathfrak{M}}f$ stets sinnvoll sei, und mit f auch Rf (falls es sinnvoll ist) stets zu $\mathfrak{M}$ gehöre. Wenn R ein H.O. ist, so ist dies auch hinreichend*[41].

Beweis. Die Notwendigkeit haben wir schon vorher erkannt, es bleibt zu zeigen, daß dies bei H.O. R auch hinreichend ist, d. h. $RP_{\mathfrak{M}}f = P_{\mathfrak{M}}Rf$ zu beweisen. Da $RP_{\mathfrak{M}}f$ nach Annahme zu $\mathfrak{M}$ gehört, ist die Zugehörigkeit von $Rf - RP_{\mathfrak{M}}f = R(f - P_{\mathfrak{M}}f)$ zu $\mathfrak{H} - \mathfrak{M}$ zu zeigen, oder da $f - P_{\mathfrak{M}}f$ zu $\mathfrak{H} - \mathfrak{M}$ gehört: wenn g zu $\mathfrak{H} - \mathfrak{M}$ gehört, so gehört auch Rg dazu (wenn es Sinn hat).

Sei Rf sinnvoll, dann gehört $RP_{\mathfrak{M}}f$ zu $\mathfrak{M}$, also ist

$$(f, P_{\mathfrak{M}}Rg) = (P_{\mathfrak{M}}f, Rg) = (RP_{\mathfrak{M}}f, g) = 0.$$

Da diese f überall dicht liegen, ist $P_{\mathfrak{M}}Rg$ zu allem orth., also $= 0$ — daher gehört Rg zu $\mathfrak{H} - \mathfrak{M}$.

Satz 21. *$\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ sei eine (eventuell abbrechende) Folge von paarweise orth. abg. lin. M., die zusammen die abg. lin. M. $\mathfrak{H}$ aufspannen. R sei ein abg. lin. Operator, der durch jede von ihnen reduziert wird, eine in $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ gelegenen Teile seien bzw. $R_1, R_2, \ldots$.*

Dann wird R durch $R_1, R_2, \ldots$ bestimmt, und zwar folgendermaßen: Rf hat dann und nur dann Sinn, wenn $f = f_1 + f_2 + \ldots$ ist ($f_1, f_2, \ldots$ aus bzw. $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$), alle $R_1f_1, R_2f_2, \ldots$ Sinn haben, beide Reihen $f_1 + f_2 + \ldots$, $R_1f_1 + R_2f_2 + \ldots$ konvergieren (wenn die Folge $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ abbricht, so fällt diese Konvergenzannahme natürlich fort), und zwar ist dann $Rf = R_1f_1 + R_2f_2 + \ldots$.

Beweis. Daß R an den angeführten Stellen Sinn hat, und zwar die angegebenen Werte annimmt, das folgt sofort aus der Definition von $R_1, R_2, \ldots$ sowie dem abg. lin. und durch $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ reduzierten Charakter von R. Wir müssen noch beweisen: wenn Rf Sinn hat, so hat f die obige Form.

Wir setzen $f_p = P_{\mathfrak{M}_p}f$ $(p = 1, 2, \ldots)$, f_p liegt dann in $\mathfrak{M}_p$, und die $P_{\mathfrak{M}_p}$ (wie die $\mathfrak{M}_p$) sind paarweise orth. Auch ist $R_pf_p = Rf_p = RP_{\mathfrak{M}_p}f = P_{\mathfrak{M}_p}Rf$ Aus der Orth. der $P_{\mathfrak{M}_p}$ folgt, daß die $P_{\mathfrak{M}_1}$, $P_{\mathfrak{M}_1} + P_{\mathfrak{M}_2}$, $P_{\mathfrak{M}_1} + P_{\mathfrak{M}_2} + P_{\mathfrak{M}_3}, \ldots$ eine (natürlich wachsende) Folge von P. O. bilden, die nach Satz 19 einen P.O. E zum Limes haben. Aus Stetigkeitsgründen sind alle Teile von E, also die $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ Teile der zu E gehörigen abg. lin. M. $\mathfrak{M}$. Nach der Annahme über die $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ ist daher $\mathfrak{M} = \mathfrak{H}$, $E = 1$.

---

[41]) In endlichvieldimensionalen Räumen ist dies auch bei unitären Operatoren hinreichend, in $\mathfrak{H}$ nicht; das ist aber hier belanglos.

Also sind die Reihen $f_1 + f_2 + \ldots$, $R_1 f_1 + R_2 f_2 + \ldots$ konvergent, und sie haben die bzw. Summen $1 f = f$, $1 R f = R f$. Damit ist alles bewiesen.

V. Die Cayleysche Transformation.

Nach diesen Vorbereitungen können wir unsere eigentliche Aufgabe in Angriff nehmen, die genauere Untersuchung der verschiedenen Arten von H. O. Die Grundlage dieser Überlegungen ist ein Kunstgriff, der eine Übertragung der Cayleyschen Transformation der Algebra auf unser Gebiet ist.

Sei R ein abg. lin. H. O. und $\mathfrak{A}$ sein Definitionsbereich. $\mathfrak{A}$ ist eine lin. M. und überall dicht (also nur dann abg., wenn es $= \mathfrak{H}$, d. h. R überall sinnvoll ist, was keineswegs vorausgesetzt wird). Wir betrachten die beiden Operatoren $R \pm i 1$, die auch in $\mathfrak{A}$ definiert sind. Man berechnet mühelos:

$$((R \pm i 1) f, (R \pm i 1) g) = (R f, R g) \pm (R f, i g) \pm (i f, R g) + (f, g)$$
$$= (R f, R g) + (f, g),$$

also insbesondere für $f = g$

$$|(R \pm i 1) f| = |(R - i 1) f| = \sqrt{|R f|^2 + |f|^2}.$$

Somit hat $f \neq 0$ $(R + i 1) f \neq 0$, $(R - i 1) f \neq 0$ zur Folge, und wegen der Lin. $f \neq g$ die Folge $(R + i 1) f \neq (R + i 1) g$, $(R - i 1) f \neq (R - i 1) g$. Wenn $\mathfrak{A}$ durch $R \pm i 1$ auf $\mathfrak{E}$ bzw. $\mathfrak{F}$ abgebildet wird, so sind die Abbildungen demnach alle ein-eindeutig. Insbesondere sei die Abbildung von $\mathfrak{E}$ auf $\mathfrak{F}$ U. Da $R \pm i 1$, wie R, abg. lin. sind, gilt dasselbe von U; weiter folgt aus der Gleichung von vorhin $|U f| = |f|$. Nach Def. 6 ist daher U ein längentreuer Operator, da $\mathfrak{E}$, $\mathfrak{F}$ sein Definitionsbereich bzw. Wertevorrat ist, sind nach Satz 10 beide abg. lin. M.

Satz 22. *Sei R ein abg. lin. H. O. und $\mathfrak{A}$ sein Definitionsbereich; die Operatoren $R \pm i 1$ mögen $\mathfrak{A}$ auf $\mathfrak{E}$ bzw. $\mathfrak{F}$ abbilden. Diese Abbildungen von $\mathfrak{A}$, $\mathfrak{E}$, $\mathfrak{F}$ aufeinander sind ein-eindeutig, und insbesondere ist die Abbildung von $\mathfrak{E}$ auf $\mathfrak{F}$ ein längentreuer Operator U. $\mathfrak{E}$, $\mathfrak{F}$ sind beide abg. lin. M.*

Beweis. Wurde soeben erbracht.

Definition 14. Die Operatoren R, U aus Satz 22 nennen wir Cayleysche Transformierte voneinander, R ist immer ein abg. lin. H. O., U immer längentreu[42]).

---

[42]) Zunächst wissen wir nur, daß jedes abg. lin. H. O. R genau eine Cayleysche Transformierte U (längentreu) hat. Der Gültigkeitsbereich der Umkehrung wird aus Satz 24 hervorgehen.

Satz 23. *Ein längentreuer Operator U ist dann und nur dann Cayleysche Transformierte des abg. lin. H. O. R, wenn*

α) $\mathfrak{A}$ die Menge aller $\varphi - U\varphi$, φ von $\mathfrak{E}$, ist,

β) für alle φ von $\mathfrak{E}$ $R(\varphi - U\varphi) = i(\varphi + U\varphi)$ ist.

Beweis. Sei U die Cayleysche Transformierte von R. Wenn f zu $\mathfrak{A}$ gehört, so gehört $(R + i1)f = \varphi$ zu $\mathfrak{E}$, es ist $U\varphi = (R - i1)f$, also $\varphi - U\varphi = 2if$, $f = \frac{\varphi}{2i} - U\frac{\varphi}{2i}$. Gehört umgekehrt φ zu $\mathfrak{E}$, so ist $(R + i1)f = \varphi$ für ein f von $\mathfrak{A}$, also $(R - i1)f = U\varphi$, $\varphi - U\varphi = 2if$, d. h. $\varphi - U\varphi$ gehört zu $\mathfrak{A}$. Und schließlich ist $\varphi + U\varphi = 2Rf$, also $R(\varphi - U\varphi) = 2iRf = i(\varphi + U\varphi)$. Also sind α), β) erfüllt.

Genüge umgekehrt U α), β). Wenn f zu $\mathfrak{A}$ gehört, so ist $f = \varphi - U\varphi$, φ von $\mathfrak{E}$ (α)), $Rf = i(\varphi + U\varphi)$ (β)), also $(R + i1)f = 2i\varphi$, $(R - i1)f = 2iU\varphi$. Somit gehört $(R + i1)f$ zu $\mathfrak{E}$, und es ist $U(R + i1)f = (R - i1)f$. Wenn wir dagegen von einem φ von $\mathfrak{E}$ ausgehen, so gehört $f = \varphi - U\varphi$ zu $\mathfrak{A}$, also ist $Rf = i(\varphi + U\varphi)$, $2i\varphi = (R + i1)f$, $\varphi = (R + i1)\frac{f}{2i}$. Somit ist $\mathfrak{E}$ die Menge aller $(R + i1)f$, f von $\mathfrak{A}$, und $U(R + i1)f = (R - i1)f$, d. h. U wirklich die Cayleysche Transformierte von R.

Satz 24. *Zum längentreuen Operator U gibt es dann und nur dann einen abg. lin. H. O. R, von dem er Cayleysche Transformierte ist, wenn die Menge aller $\varphi - U\varphi$, φ von $\mathfrak{E}$ ($\mathfrak{E}$ ist der Definitionsbereich von U), überall dicht ist. Und zwar ist dann R durch U eindeutig bestimmt.*

Beweis. Die Notwendigkeit ist klar nach α), Satz 23, da $\mathfrak{A}$ überall dicht ist. Die Hinreichendheit sowie die eindeutige Bestimmtheit von R durch U (nach α), β) in Satz 23) steht gleichfalls fest, sobald erkannt ist, daß β) für R keine Unmöglichkeit involviert, und daß das so definierte R ein abg. lin. H. O. ist.

Das erstere folgt daraus, daß für $\varphi \neq \psi$ $\varphi - U\varphi \neq \psi - U\psi$ ist, d. h. für $\varphi \neq 0$ $\varphi - U\varphi \neq 0$. In der Tat hat $\varphi = U\varphi$ die Folge, daß für alle χ von $\mathfrak{E}$

$$(\varphi, \chi) = (U\varphi, U\chi) = (\varphi, U\chi), \quad (\varphi, \chi - U\chi) = 0$$

ist, da aber die $\chi - U\chi$ überall dicht liegen, ist φ zu einer überall dichten Menge orth., also zu ganz $\mathfrak{H}$, also $\varphi = 0$. Das letztere beweisen wir so. Die abg. Lin. von R folgt aus der von U, die überall Dichtigkeit des Definitionsbereiches von R (d. h. der $\varphi - U\varphi$) wurde vorausgesetzt — es bleibt noch übrig $(f, Rg) = (Rf, g)$ zu beweisen, d. h. daß (f, Rg) beim Vertauschen von f und g in seine komplex-Konjugierte übergeht. Sei also $f = \varphi - U\varphi$, $g = \psi - U\psi$, dann ist

$$
\begin{aligned}
(f, Rg) &= (\varphi - U\varphi, i(\psi + U\psi)) \\
&= -i\{(\varphi, \psi) - (U\varphi, \psi) + (\varphi, U\psi) - (U\varphi, U\psi)\} \\
&= i((U\varphi, \psi) - \overline{(U\psi, \varphi)}) = i(U\varphi, \psi) + \overline{i(U\psi, \varphi)},
\end{aligned}
$$

was die Behauptung in Evidenz setzt.

Satz 25. *R, S seien abg. lin. H.O., U, V ihre Cayleyschen Transformierten. R ist dann und nur dann Forts. oder eig. Forts. von S, wenn U Forts. bzw. eig. Forts. von V ist. — $\mathfrak{M}$ sei eine abg. lin. M., es reduziert R dann und nur dann, wenn es U reduziert.*

Beweis. Man schließt jeweils mühelos mit Hilfe von Satz 22 von R auf U, und mit Hilfe von Satz 23 von U auf R.

In den vorhergehenden Sätzen haben wir die Beurteilung der Fortsetzbarkeitsverhältnisse von R auf die viel einfacheren von U zurückgeführt. Die Cayleyschen Transformierten aller R sind alle U des Satzes 5, und dabei ist zu beachten: wenn R Forts. von S sein soll, also U Forts. eines Satz 5 genügenden V, so braucht es nur längentreu zu sein[43]). An Stelle der zunächst recht unübersichtlichen Forts. von R brauchen wir also bloß die einfach angebbaren längentreuen Forts. von V (das selbst längentreu ist) zu betrachten. Um uns über derartige Fragen zu orientieren, beweisen wir zwei einfache (anschaulich-geometrische) Sätze über längentreue Abbildungen und abg. lin. M.

Zunächst bemerken wir: unter der Dimensionszahl einer abg. lin. M. $\mathfrak{M}$ verstehen wir die obere Grenze aller (endlichen!) n, für die es in $\mathfrak{M}$ n lin. unabh. Elemente gibt — also eine der Zahlen $0, 1, 2, \ldots, \infty$. (Eine wesentlich allgemeinere Definition der Dimensionszahl von Mengen in $\mathfrak{H}$ werden wir im nächsten Kapitel, Def. 16, geben.)

Satz 26. *$\mathfrak{M}$ sei eine abg. lin. M., $\varphi_1, \ldots, \varphi_n$ ein norm. orth. System, welches diese abg. lin. M. aufspannt (es gibt gewiß solche, vgl. den Beweis von Satz 13; es ist $n = 0, 1, 2, \ldots, \infty$, im letzteren Falle bricht die Folge $\varphi_1, \varphi_2, \ldots$ nicht ab), dann ist $\mathfrak{M}$ n-dimensional.*

Zwei abg. lin. M. $\mathfrak{M}$, $\mathfrak{N}$ können dann und nur dann Definitionsbereich bzw. Wertevorrat eines längentreuen Operators U sein, wenn sie gleichvieldimensional sind. Sei $\varphi_1, \ldots, \varphi_n$ ein $\mathfrak{M}$ aufspannendes norm. orth. System, dann entsprechen die $\mathfrak{N}$ aufspannenden norm. orth. Systeme $\psi_1, \ldots, \psi_n$ diesen U ein-eindeutig: jedem wird durch die Bedingungen $U\varphi_1 = \psi_1, \ldots, U\varphi_n = \psi_n$ genau ein solches U zugeordnet.

Beweis. $\varphi_1, \ldots, \varphi_n$ seien ein norm. orth. System, das die abg. lin. M. $\mathfrak{M}$ aufspannt. Da $\varphi_1, \ldots, \varphi_n$ lin. unabh. sind, hat $\mathfrak{M}$ $\geq n$ Dimensionen,

---

[43]) Denn die $\varphi - U\varphi$ umfassen die $\varphi - V\varphi$, liegen also wie diese überall dicht.

für $n = \infty$ sind wir also fertig, für endliches n ist zu zeigen: $n+1$ Elemente von $\mathfrak{M}$ sind nie lin. unabh. In der Tat müssen sie die Form

$$f_\mu = a_{\mu 1}\varphi_1 + \ldots + a_{\mu n}\varphi_n \qquad (\mu = 1, \ldots, n+1)$$

haben, und zwischen den $n+1$ n-dimensionalen Vektoren

$$a_{\mu 1}, \ldots, a_{\mu n} \qquad (\mu = 1, \ldots, n+1)$$

muß eine lineare Relation bestehen, also auch zwischen den f_μ.

Gehen wir zur zweiten und dritten Behauptung über. Sei $\mathfrak{M}$ der Definitionsbereich und $\mathfrak{N}$ der Wertevorrat des längentreuen Operators U. Wenn $\varphi_1, \ldots, \varphi_n$ ein $\mathfrak{M}$ aufspannendes norm. orth. System ist, so ist (wegen der Längentreue von U) $\psi_1 = U\varphi_1, \ldots, \psi_n = U\varphi_n$ ein $\mathfrak{N}$ aufspannendes norm. orth. System. Und $\mathfrak{M}$ wie $\mathfrak{N}$ haben n Dimensionen.

Umgekehrt seien $\mathfrak{M}$, $\mathfrak{N}$ zwei gleichviel- (etwa n-) dimensionale abg. lin. M. Sei $\varphi_1, \ldots, \varphi_p$ bzw. $\psi_1, \ldots, \psi_q$ je ein $\mathfrak{M}$ bzw. $\mathfrak{N}$ aufspannendes norm. orth. System, es muß $p = q = n$ sein. Die Reihen $\sum_{r=1}^{n} x_r \varphi_r$ und $\sum_{r=1}^{n} x_r \psi_r$ konvergieren unter denselben Umständen: bei endlichem n immer, bei $n = \infty$, falls $\sum_{r=1}^{\infty} |x_r|^2$ endlich ist (Satz 5) — also allenfalls, wenn $\sum_{r=1}^{\infty} |x_r|^2$ endlich ist; sie erschöpfen also $\mathfrak{M}$ bzw. $\mathfrak{N}$ ($\sum_{r=1}^{n} x_r \varphi_r$ hat mit φ_s das innere Produkt x_s, es bestimmt also $x_1, \ldots, x_n$ eindeutig). Wir können also einen Operator U durch $U\left(\sum_{r=1}^{n} x_r \varphi_r\right) = \sum_{r=1}^{n} x_r \psi_r$ definieren, er hat den Definitionsbereich $\mathfrak{M}$ und den Wertevorrat $\mathfrak{N}$, und ist offenbar linear. Es ist $\left(f = \sum_{r=1}^{n} x_r \varphi_r\right)$

$$|f|^2 = |Uf|^2 = \sum_{r=1}^{n} |x_r|^2, \qquad |f| = |Uf|,$$

also der Operator stetig, und weil sein Definitionsbereich abg. ist, ist er auch selbst abg. Also ist U längentreu.

Dabei hat Längentreue, also insbesondere abg. Lin., und $U\varphi_\mu = \psi_\mu$ $(\mu = 1, \ldots, n)$ $U\left(\sum_{r=1}^{n} x_r \varphi_r\right) = \sum_{r=1}^{n} x_r \psi_r$ zur Folge; daher ist unser U durch diese Eigenschaften eindeutig festgelegt. Damit sind alle Behauptungen unseres Satzes bewiesen.

Satz 27. *U sei ein längentreuer Operator, $\mathfrak{E}$, $\mathfrak{F}$ sein Definitionsbereich bzw. Wertevorrat (also zwei abg. lin. M.). Wir gewinnen alle längentreuen Forts. von U auf die folgende Weise:*

Man wähle zwei gleichvieldimensionale abg. lin. M. $\mathfrak{M}$, $\mathfrak{N}$, Teile von $\mathfrak{H} - \mathfrak{E}$, $\mathfrak{H} - \mathfrak{F}$, und einen längentreuen Operator U' mit diesem Definitionsbereich bzw. Wertevorrat. Diese bestimmen genau eine längentreue

Forts. V von U, die so charakterisiert ist: Definitionsbereich bzw. Wertevorrat sind $\mathfrak{E}+\mathfrak{M}$ bzw. $\mathfrak{F}+\mathfrak{N}$, in $\mathfrak{E}$ gilt $Vf=Uf$, in $\mathfrak{M}$ gilt $Vf=U'f$. — Indem man $\mathfrak{M}, \mathfrak{N}, U'$ auf alle möglichen Weisen wählt, gewinnt man alle längentreuen Forts. V von U, und zwar jede genau einmal.

Beweis. Daß jede längentreue Forts. V von U so entsteht, ist klar: seien $\mathfrak{K}$, $\mathfrak{L}$ Definitionsbereich bzw. Wertevorrat von V, so genügt es, $\mathfrak{M}=\mathfrak{K}-\mathfrak{E}$, $\mathfrak{N}=\mathfrak{L}-\mathfrak{F}$, $U'=$ das auf $\mathfrak{M}$ eingeschränkte V zu setzen, und alles ist erfüllt. Auch daß unsere Bedingungen genau ein V (wenn überhaupt etwas) festlegen, ist klar: genügen ihr V' und V'', so ist $V'f=V''f$ in ganz $\mathfrak{E}$ und ganz $\mathfrak{M}$, also wegen der Lin. auch im ganzen Definitionsbereiche $\mathfrak{E}+\mathfrak{M}$ — d. h. $V'=V''$. Es bleibt daher nur übrig zu zeigen, daß wirklich jedes System $\mathfrak{M}, \mathfrak{N}, U'$ mindestens ein längentreues V bestimmt.

Da $\mathfrak{M}$ bzw. $\mathfrak{N}$ Teil von $\mathfrak{H}-\mathfrak{E}$ bzw. $\mathfrak{H}-\mathfrak{F}$, also orth. zu $\mathfrak{E}$ bzw. $\mathfrak{F}$ ist, sind $\mathfrak{M}$, $\mathfrak{E}$ lin. unabh., also in $\mathfrak{E}+\mathfrak{M}$ die Zerlegung $f=g+h$, g von $\mathfrak{E}$, h von $\mathfrak{M}$, eindeutig. Wir definieren also $V(g+h)=Ug+U'h$ (g von $\mathfrak{E}$, h von $\mathfrak{M}$). Da g, h orth. sind, und ebenso $Ug, U'h$, und U, U' längentreu, so ist $|V(g+h)|^2=|Ug+U'h|^2=|Ug|^2+|U'h|^2 =|g|^2+|h|^2$, $|g+h|^2=|g|^2+|h|^2$. V ist in der abg. lin. M. $\mathfrak{E}+\mathfrak{M}$ definiert, offenbar lin., nach dem vorigen stetig — also abg., und nach dem vorigen längentreu. Außerdem, wie gewünscht, Forts. von U und U'.

Durch die Sätze 26, 27 übersehen wir alle längentreuen Forts. eines längentreuen U, also auch alle abg. lin. H. O. Forts. eines abg. lin. H. O. R; wir können sie alle effektiv konstruieren. So sehen wir z. B.: R ist max., d. h. es hat keine eig. Forts., wenn U keine hat — also wenn es unmöglich ist in Satz 26 $\mathfrak{M}$, $\mathfrak{N}$ von (0) verschieden (vgl. Anm. [35])) zu wählen. Dies ist offenbar nur für $\mathfrak{H}-\mathfrak{E}=(0)$ oder $\mathfrak{H}-\mathfrak{F}=(0)$ der Fall, d. h. für $\mathfrak{E}=\mathfrak{H}$ oder $\mathfrak{F}=\mathfrak{H}$. Die genauere Untersuchung dieser Fragen soll in den drei nächsten Kapiteln durchgeführt werden.

VI. Erweiterungselemente.

Wir wollen die bereits im letzten Kapitel zur Geltung gekommenen abg. lin. M. $\mathfrak{E}$, $\mathfrak{F}$ in ihrem Verhältnis zu den Begriffsbildungen von Def. 8 näher charakterisieren.

Für die Zwecke dieses Kapitels ist es bequem, die Terminologie des Restklassen-Kalküls der Algebra zu verwenden: wenn $\mathfrak{M}$ eine (nicht notwendig abg.!) lin. M. ist, so bedeute $f=g \ldots \mathrm{mod}\, \mathfrak{M}$, daß $f-g$ zu $\mathfrak{M}$ gehört. $R, U, \mathfrak{A}, \mathfrak{E}, \mathfrak{F}$ mögen die bisherige Bedeutung haben, wie z. B. in Satz 22.

Satz 28. $\mathfrak{H} - \mathfrak{E}$ *bzw.* $\mathfrak{H} - \mathfrak{F}$ *ist die Menge aller erweiterbaren* f, *denen* R if *bzw.* $-if$ *zuordnet.*

Beweis. Daß f^* dem f zugeordnet ist, bedeutet nach Satz 23

$$(f^*, \varphi - U\varphi) = (f, i(\varphi + U\varphi)) = (-if, \varphi + U\varphi).$$

$f^* = if$ bzw. $= -if$ bedeutet also, daß f (und mit ihm $f^* = \pm if$) auf alle φ (von $\mathfrak{E}$) bzw. alle $U\varphi$ (d. h. ganz $\mathfrak{F}$) orth. steht; d. h. daß es zu $\mathfrak{H} - \mathfrak{E}$ bzw. $\mathfrak{H} - \mathfrak{F}$ gehört. –

Daher gehört ganz $\mathfrak{H} - \mathfrak{E}$ bzw. ganz $\mathfrak{H} - \mathfrak{F}$ (mit Ausnahme der 0) zur $+$- bzw. $-$-Klasse der Erweiterungselemente.

Satz 29. $\mathfrak{A} + (\mathfrak{H} - \mathfrak{E}) + (\mathfrak{H} - \mathfrak{F})$ *ist die Menge aller Erweiterungselemente von* R.

Beweis. Die Erweiterungselemente bilden offenbar eine lin. M., da diese, wie wir bereits wissen, $\mathfrak{A}$, $\mathfrak{H} - \mathfrak{E}$, $\mathfrak{H} - \mathfrak{F}$ umfassen muß, ist $\mathfrak{A} + (\mathfrak{H} - \mathfrak{E}) + (\mathfrak{H} - \mathfrak{F})$ Teil von ihr; es bleibt zu zeigen, daß sie nicht größer ist.

Sei also f Erweiterungselement, nach dem obigen bedeutet das:

$$(f^*, \varphi - U\varphi) = (-if, \varphi + U\varphi), \quad (f^* + if, \varphi) = (f^* - if, U\varphi)$$

(für alle φ von $\mathfrak{E}$, f^* fest). Sei U^{-1} die Inverse von U, mit $\mathfrak{F}$, $\mathfrak{E}$ als Definitionsbereich bzw. Wertevorrat. Es ist daher $P_{\mathfrak{F}} U = U$, $U^{-1} U = 1$ und $P_{\mathfrak{E}} U^{-1} = U^{-1}$, $U U^{-1} = 1$, soweit diese Operatoren Sinn haben, d. h. in $\mathfrak{E}$ bzw. $\mathfrak{F}$. Nunmehr ist

$$(f^* - if, U\varphi) = (f^* - if, P_{\mathfrak{F}} U\varphi) = (P_{\mathfrak{F}} f^* - i P_{\mathfrak{F}} f, U\varphi)$$
$$= (U U^{-1} P_{\mathfrak{F}} f^* - i U U^{-1} P_{\mathfrak{F}} f, U\varphi) = (U^{-1} P_{\mathfrak{F}} f^* - i U^{-1} P_{\mathfrak{F}} f, \varphi),$$

also

$$(\{f^* + if\} - \{U^{-1} P_{\mathfrak{F}} f^* - i U^{-1} P_{\mathfrak{F}} f\}, \varphi) = 0.$$

Da dies für alle φ von $\mathfrak{E}$ gilt, so ist

$$P_{\mathfrak{E}}(\{f^* + if\} - \{U^{-1} P_{\mathfrak{F}} f^* - i U^{-1} P_{\mathfrak{F}} f\}) = 0,$$
$$\{P_{\mathfrak{E}} f^* + i P_{\mathfrak{E}} f\} - \{U^{-1} P_{\mathfrak{F}} f^* - i U^{-1} P_{\mathfrak{F}} f\} = 0,$$
$$- U^{-1} P_{\mathfrak{F}} f^* + P_{\mathfrak{E}} f^* = - i (U^{-1} P_{\mathfrak{F}} f + P_{\mathfrak{E}} f),$$

oder nach Anwendung von $-U$

$$P_{\mathfrak{F}} f^* - U P_{\mathfrak{E}} f^* = i (P_{\mathfrak{F}} f + U P_{\mathfrak{E}} f).$$

Weil allgemein $g - P_{\mathfrak{M}} g = P_{\mathfrak{H}-\mathfrak{M}} g = 0 \ldots \bmod (\mathfrak{H} - \mathfrak{M})$ gilt, ist $P_{\mathfrak{E}} g = P_{\mathfrak{F}} g \ldots \bmod ((\mathfrak{H} - \mathfrak{E}) \dotplus (\mathfrak{H} - \mathfrak{F}))$. Darum folgt aus der obigen Gleichung

$$P_{\mathfrak{E}} f^* - U P_{\mathfrak{E}} f^* = i (P_{\mathfrak{E}} f + U P_{\mathfrak{E}} f) \ldots \bmod ((\mathfrak{H} - \mathfrak{E}) \dotplus (\mathfrak{H} - \mathfrak{F})).$$

Die linke Seite gehört ihrer Natur nach zu $\mathfrak{A}$, also gehört die rechte zu $\mathfrak{A}+(\mathfrak{H}-\mathfrak{E})+(\mathfrak{H}-\mathfrak{F})$; ebenso gehört $P_{\mathfrak{E}}f - UP_{\mathfrak{E}}f$ seiner Natur nach zu $\mathfrak{A}$, aus beidem folgt, daß auch $P_{\mathfrak{E}}f$ zu $\mathfrak{A}+(\mathfrak{H}-\mathfrak{E})+(\mathfrak{H}-\mathfrak{F})$ gehört. Da aber $f = P_{\mathfrak{E}}f \ldots \bmod (\mathfrak{H}-\mathfrak{E})$ ist, gilt dasselbe von f — womit alles bewiesen ist.

Satz 30. *$\mathfrak{A}$, $\mathfrak{H}-\mathfrak{E}$, $\mathfrak{H}-\mathfrak{F}$ sind lin. unabh.*

Beweis. Wir müssen aus $f+\varphi+\psi=0$, f von $\mathfrak{A}$, φ von $\mathfrak{H}-\mathfrak{E}$, ψ von $\mathfrak{H}-\mathfrak{F}$, folgern, daß $f=\varphi=\psi=0$ ist, oder auch: daraus, daß φ zu $\mathfrak{H}-\mathfrak{E}$, ψ zu $\mathfrak{H}-\mathfrak{F}$, $\varphi+\psi$ zu $\mathfrak{A}$ gehört, $\varphi=\psi=0$. Sei also diese letztere Prämisse erfüllt.

φ, ψ, $\varphi+\psi$ sind nach früher Gesagtem Erweiterungselemente, denen R $i\varphi$, $-i\psi$, $R(\varphi+\psi)$ (dieses hat Sinn!) zuordnet. Da $\varphi+\psi$ offenbar auch $i(\varphi-\psi)$ zugeordnet wird, muß dieses $=R(\varphi+\psi)$ sein. Weil φ, ψ $i\varphi$ bzw. $-i\psi$ zugeordnet sind, so muß

$$(i\varphi, \varphi+\psi) = (\varphi, R(\varphi+\psi)) = (\varphi, i(\varphi-\psi)), \quad 2i(\varphi,\varphi)=0, \ \varphi=0,$$

$$(-i\psi, \varphi+\psi) = (\psi, R(\varphi+\psi)) = (\psi, i(\varphi-\psi)), \quad -2i(\psi,\psi)=0, \ \psi=0$$

gelten. —

Wir können also jedes Erweiterungselement f auf eine und nur eine Weise in $f=f^0+f^++f^-$, f^0 von $\mathfrak{A}$, f^+ von $\mathfrak{H}-\mathfrak{E}$, f^- von $\mathfrak{H}-\mathfrak{F}$, zerlegen, und umgekehrt ist jedes solche f ein Erweiterungselement. Wir nennen f^0, f^+, f^- bzw. die 0-, +-, —-Komponente von f. Da f^0, f^+, f^- Rf^0, if^+, $-if^-$ zugeordnet wird, so ordnet R dem $f=f^0+f^++f^-$ $f^* = Rf_0+if^+-if^-$ zu. — Diese Zerlegung ist von Wichtigkeit, weil sie eine nähere Kennzeichnung der 0-, +-, —-Klassen erlaubt:

Satz 31. *Ein Erweiterungselement f gehört zur 0-, +-, —-Klasse, je nachdem ob $|f^+| =, >, < |f^-|$ ist.*

Beweis. Wegen $f=f^0+f^++f^-$, $f^*=Rf^0+if^+-if^-$ und der Zugehörigkeit von if^+, $-if^-$ zu f^+ bzw. f^- ist:

$$\begin{aligned}(f^*, f) &= (Rf^0+if^+-if^-, f^0+f^++f^-)\\ &= (Rf^0, f^0) + \{(Rf^0, f^+)+(if^+, f^0)\} + \{(Rf^0, f^-)+(-if^-, f_0)\}\\ &\quad + \{(if^+, f^-)+(-if^-, f^+)\} + \{(if^+, f^+)+(-if^-, f^-)\}\\ &= (f^0, Rf^0) + \{(f^0, if^+)+(if^+, f^0)\} + \{(f^0, -if^-)+(-if^-, f_0)\}\\ &\quad + \{(f^+, -if^-)+(-if^-, f^+)\} + i\{|f^+|^2-|f^-|^2\}.\end{aligned}$$

Das erste Glied und die drei ersten $\{\ \}$ sind offenbar reell, die letzte $\{\ \}$ offenbar rein imaginär, woraus

$$\mathfrak{J}(f^*, f) = |f^+|^2 - |f^-|^2$$

folgt — also die Behauptung.

VII. Die Defektindizes.

Definition 15. $R, U, \mathfrak{A}, \mathfrak{E}, \mathfrak{F}$ mögen die bisherige Bedeutung haben, wie z. B. in Satz 22. m, n seien die Dimensionszahlen von $\mathfrak{H} - \mathfrak{E}$ bzw. $\mathfrak{H} - \mathfrak{F}$.

Wir nennen m, n die Defektindizes von R. (Es ist $m, n = 0, 1, 2, \ldots, \infty$).

Diese Zahlen werden bei den Fragen der Max. und Hypermax. eine entscheidende Rolle spielen. Zwischen ihnen und den drei Klassen der Erweiterungselemente besteht ein enger Zusammenhang, den wir nun dartun werden. Zunächst verallgemeinern wir den Begriff der Dimensionszahl in $\mathfrak{H}$.

Definition 16. $\mathfrak{M}$ sei eine (nicht notwendig abg.!) lin. M., $\mathfrak{U}$ eine Teilmenge von $\mathfrak{H}$ (die nicht einmal lin. M. zu sein braucht). Unter der Dimensionszahl von $\mathfrak{U} \ldots \mathrm{mod}\ \mathfrak{M}$ verstehen wir die obere Grenze aller (endlichen!) n, für die es n lin. unabh. Elemente gibt, die zusammen mit ihrer lin. M.[44]) (mit eventueller Ausnahme der 0[45])) ganz in $\mathfrak{U}$ liegen und außerhalb von $\mathfrak{M}$. n ist also eine der Zahlen $0, 1, 2, \ldots, \infty$.

Übrigens wird im folgenden $\mathfrak{U}$ doch nicht ganz willkürlich sein, es wird vielmehr stets die folgenden Eigenschaften haben: Erstens ist es eine Restklassenmenge mod $\mathfrak{M}$, d. h. von zwei f, g, $f = g \ldots \mathrm{mod}\ \mathfrak{M}$, gehört keins oder beide zu $\mathfrak{U}$, und zweitens gehören mit f auch alle af $(a \neq 0)$ zu $\mathfrak{U}$. Man sieht sofort, daß unter diesen $\mathfrak{U}$ die leere Menge und $\mathfrak{M}$ die einzigen sind, die die Dimensionszahl 0 haben. Weiter ist es klar, daß die Dimensionszahl einer abg. lin. M. $\mathfrak{U} \ldots \mathrm{mod}\ (0)$ mit der von Kap. VI übereinstimmt.

Die $+$-, $-$- und 0-Klassen sind solche Mengen $\mathfrak{U}$ mod $\mathfrak{A}$, ebenso die Vereinigungsmengen irgendwelcher von ihnen; da die zwei ersten 0 nicht enthalten, wohl aber die dritte, so müssen sie im Falle der 0-Dimensionalität leer bzw. gleich $\mathfrak{A}$ sein.

Satz 32. *Die $+$-Klasse, sowie ihre Vereinigung mit der* 0-*Klasse, ist* mod $\mathfrak{A}$ *m-dimensional; die $-$-Klasse, sowie ihre Vereinigung mit der* 0-*Klasse ist* mod $\mathfrak{A}$ *n-dimensional; die* 0-*Klasse ist* mod $\mathfrak{A}$ Min (m, n)-*dimensional.*

Beweis. Sei $p \leqq m$ und endlich (es ist ja eventuell $m = \infty$), und $f_1, \ldots, f_p$ lin. unabh. Elemente von $\mathfrak{H} - \mathfrak{E}$ (das ja m-dimensional ist). Die ganze lin. M. derselben liegt in $\mathfrak{H} - \mathfrak{E}$, also, mit Ausnahme der 0, in der $+$-Klasse — und somit von selbst außerhalb $\mathfrak{A}$. Also hat die $+$-Klasse $\geqq m$-Dimensionen mod $\mathfrak{A}$, wenn wir noch zeigen, daß ihre Vereinigung

---

[44]) Die wegen der Endlichkeit von n von selbst abg. ist.

[45]) Diese Ausnahme wird sich im folgenden als praktisch erweisen.

mit der 0-Klasse $\leqq m$ hat, sind die zwei ersten Behauptungen bewiesen. Für $m = \infty$ ist dies trivial, bei endlichem m ist zu zeigen: wenn $f_1, \ldots, f_{m+1}$ lin. unabh. Erweiterungselemente sind, so liegt gewiß ein

$$a_1 f_1 + \ldots + a_{m+1} f_{m+1} \neq 0$$

in $\mathfrak{A}$ oder in der $-$-Klasse.

Sei $f_\mu = f_\mu^0 + f_\mu^+ + f_\mu^-$ $(\mu = 1, \ldots, m+1)$, die f_μ^+ liegen alle in dem m-dimensionalen $\mathfrak{H} - \mathfrak{E}$, sind also nicht lin. unabh. Sei etwa

$$a_1 f_1^+ + \ldots + a_{m+1} f_{m+1}^+ = 0$$

(ohne $a_1 = \ldots = a_{m+1} = 0$, also $a_1 f_1 + \ldots + a_{m+1} f_{m+1} \neq 0$!), dann ist $a_1 f + \ldots + a_{m+1} f_{m+1} = g^0 + g^-$, g^0 von $\mathfrak{A}$, g^- von $\mathfrak{H} - \mathfrak{F}$. Für $g^- = 0$ gehört es also zu $\mathfrak{A}$, für $g^- \neq 0$ (nach Satz 31) zur $-$-Klasse.

Damit sind die zwei ersten Behauptungen bewiesen, die zwei folgenden ergeben sich ebenso durch Vertauschen von $+$ und $-$. Es bleibt die letzte. Da die Dimensionszahl der 0-Klasse mod $\mathfrak{A}$ allenfalls $\leqq m$ und $\leqq n$ ist, genügt es zu zeigen, daß sie $\geqq$ Min. (m, n) ist: d. h. zu jedem $p \leqq$ Min. (m, n) solche lin. unabh. $f_1, \ldots, f_p$ anzugeben, daß jedes $a_1 f_1 + \ldots + a_p f_p \neq 0$ zur 0-Klasse gehört, aber nicht zu $\mathfrak{A}$.

In $\mathfrak{H} - \mathfrak{E}$ gibt es gewiß p lin. unabh. Elemente, die wir gleich „orthogonalisieren" (vgl. Satz 8), d. h norm. orth. annehmen können: $\varphi_1, \ldots, \varphi_p$; ebenso seien $\psi_1, \ldots, \psi_p$ norm. orth. aus $\mathfrak{H} - \mathfrak{F}$. Wir setzen

$$f_1 = \varphi_1 + \psi_1, \quad \ldots, \quad f_p = \varphi_p + \psi_p,$$

dann ist (wenn nicht $a_1 = \ldots = a_p = 0$)

$$a_1 f_1 + \ldots + a_p f_p = (a_1 \varphi_1 + \ldots + a_p \varphi_p) + (a_1 \psi_1 + \ldots + a_p \psi_p)$$

(nach Satz 30) sicher nicht in $\mathfrak{A}$, also insbesondere $\neq 0$; und wegen

$$|a_1 \varphi_1 + \ldots + a_p \varphi_p|^2 = |a_1 \psi_1 + \ldots + a_p \psi_p|^2 = |a_1|^2 + \ldots + |a_p|^2$$

gehört es (nach Satz 31) zur 0-Klasse. Die $f_1, \ldots, f_p$ sind also lin. unabh. und leisten alles was wir wünschen.

Satz 33. *R ist dann und nur dann max., wenn $m = 0$ oder $n = 0$ ist, und hypermax., wenn $m = n = 0$ ist. Oder auch: wenn $\mathfrak{E} = \mathfrak{H}$ oder $\mathfrak{F} = \mathfrak{H}$ ist bzw. wenn $\mathfrak{E} = \mathfrak{F} = \mathfrak{H}$ ist.*

Beweis. Nach Satz 11 bedeutet die Max.: 0-Klasse gleich $\mathfrak{A}$, nach Definition 9 die Hypermax.: außerdem noch die $+$- und die $-$-Klasse leer. Nach der Bemerkung von Satz 32 bedeutet ersteres: 0-Klasse 0-dimensional, und letzteres: $+$- und $-$-Klasse 0-dimensional. Nach Satz 32 schließlich heißt dies: Min $(m, n) = 0$, d. h. $m = 0$ oder $n = 0$ bzw. $m = 0$ und $n = 0$.

Die zweite Formulierung folgt sofort aus der ersten. —

Der Gegensatz von Max. und Hypermax. tritt nunmehr klar hervor, ebenso die Rolle von m, n. Zu Satz 32 sei noch bemerkt: Wenn wir R durch $aR + b1$ ersetzen (a, b reell, $a \neq 0$; dies ist wieder ein abg. lin. H. O. mit demselben Definitionsbereich wie R), so ändern sich $\mathfrak{E}, \mathfrak{F}$ auf unübersichtliche Weise, und wir sehen daher zunächst nicht, was mit m, n geschieht. Aber $\mathfrak{A}$, die Menge der Erweiterungselemente, sowie die 0-, +-, −-Klassen ändern sich bei $a > 0$ offenbar überhaupt nicht, und bei $a < 0$ werden nur die zwei letzteren vertauscht. Daher hat Satz 32 die Konsequenz, daß sich m, n auch nicht ändern bzw. nur vertauscht werden[46].

VIII. Der Fortsetzungsprozeß.

Bereits am Schlusse des Kap. VI hatten wir eine Übersicht über alle möglichen Forts. eines abg. lin. H. O. R gewonnen; wir können jetzt entscheiden, wann und wie eine Forts. zu max. bzw. hypermax. Operatoren gelingt.

Satz 34. *Ein abg. lin. H. O. R mit den Defektindizes m, n kann dann und nur dann zu einem mit den Defektindizes m', n' fortgesetzt werden wenn es ein $p\,(= 0, 1, 2, \ldots, \infty)$ mit $m' + p = m$, $n' + p = n$ gibt.*

Beweis. Wenn wir den Fortsetzungsprozeß nach Satz 27 betrachten ($U, \mathfrak{A}, \mathfrak{E}, \mathfrak{F}$ mögen den üblichen Sinn haben, $\mathfrak{M}, \mathfrak{N}$ wie in Satz 27), so sehen wir: $\mathfrak{H} - \mathfrak{E}, \mathfrak{H} - \mathfrak{F}$ sind m- bzw. n-dimensional, und die Frage ist, ob es zwei gleichviel- (etwa p-) dimensionale abg. lin. M. $\mathfrak{M}, \mathfrak{N}$ gibt, die

---

[46]) Sei $\overline{\overline{R}}$ der folgende Operator: $\overline{\overline{R}}f$ hat Sinn für alle Erweiterungselemente f von R, und zwar ist es dann dem zugeordneten f^* gleich. $\overline{\overline{R}}$ ist offenbar abg. lin. und Forts. von R. (Wenn R Hypermax. ist, so ist offenbar $\overline{\overline{R}} = R$, andernfalls ist es nicht einmal ein H. O., denn wegen $m > 0$ oder $n > 0$ ist dann die +- oder die −-Klasse nicht leer und in diesen ist $(\overline{\overline{R}}f, f) = (f^*, f)$ nicht reell.) Nach Satz 28 sind $\mathfrak{H} - \mathfrak{E}$ und $\mathfrak{H} - \mathfrak{F}$ die Menge der Lösungen von $\overline{\overline{R}}f = if$ bzw. $= -if$. Also ist für diese Gleichungen m bzw. n die Höchstzahl von lin. unabh. Lösungen.

Nach dem vorhin Gesagten ist aber die Höchstzahl der lin. unabh. Lösungen von $a\overline{\overline{R}}f + bf = if$ gleich m bzw. n für $a >$ bzw. < 0. Die obige Gleichung besagt: $\overline{\overline{R}}f = \varrho f$, $\varrho = -\frac{b}{a} + \frac{1}{a} i$. Wir können also sagen: Wenn ϱ eine nicht reelle Zahl ist, so hat die Gleichung

$$\overline{\overline{R}}f = \varrho f$$

die Höchstzahl m bzw. n von lin. unabh. Lösungen, wenn $\mathfrak{J}\varrho >$ bzw. < 0 ist. (Bei reellem ϱ ist die Höchstzahl bekanntlich die „Vielfachheit des Punkteigenwertes" ϱ. $Rf = \varrho f$ muß bei nicht reellem ϱ unlösbar sein — außer durch $f = 0$; da sonst $(Rf, f) = \varrho(f, f)$ nicht reell ist.)

Diese Tatsachen stehen in Beziehung zu gewissen Resultaten von Carleman, vgl. das Zitat von Anm. [4]).

Teile von denselben sind, und zwar so, daß $\mathfrak{H}-\mathfrak{E}-\mathfrak{M}$, $\mathfrak{H}-\mathfrak{F}-\mathfrak{N}$ m'- bzw. n'-dimensional ist.

Allgemeiner ist die Frage diese: Wann hat eine k-dimensionale abg. lin. M. $\mathfrak{K}$ ein p-dimensionales Teil (eine abg. lin. M.) $\mathfrak{P}$, so daß $\mathfrak{K}-\mathfrak{P}$ k' Dimensionen hat? Wir werden sehen: wenn $k=k'+p$ ist. Wenn wir hierin für $\mathfrak{K}, \mathfrak{P}, k, k'$ $\mathfrak{H}-\mathfrak{E}, \mathfrak{M}, m, m'$ und $\mathfrak{H}-\mathfrak{F}, \mathfrak{N}, n, n'$ einsetzen, haben wir unseren Satz.

Notwendig ist die Bedingung: denn seien $\mathfrak{P}, \mathfrak{K}-\mathfrak{P}$ p- bzw. k'-dimensional, dann werden sie (als abg. lin. M.) von den norm. orth. Systemen $\varphi_1, \ldots, \varphi_p$ bzw. $\psi_1, \ldots, \psi_k$ aufgespannt, also $\mathfrak{K}=\mathfrak{P}+(\mathfrak{K}-\mathfrak{P})$ von $\varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_k$ (diese sind orth., weil es $\mathfrak{P}, \mathfrak{K}-\mathfrak{P}$ sind), d. h. es ist $p+k'$-dimensional, also $k=k'+p$. (Hierbei kann auch k' oder p $=\infty$ sein!) — Hinreichend ist sie ebenfalls, denn sei $\mathfrak{K}$ $k=k'+p$-dimensional, dann wird es von einem norm. orth. System $\varphi_1, \ldots, \varphi_p, \psi_1, \ldots, \psi_k$ aufgespannt (wieder darf k' oder $p=\infty$ sein!). Die von $\varphi_1, \ldots, \varphi_p$ aufgespannte abg. lin. M. sei $\mathfrak{P}$, dann wird $\mathfrak{K}-\mathfrak{P}$ von $\psi_1, \ldots, \psi_k$ aufgespannt, und wir sind am Ziele.

Satz 35. *R sei ein abg. lin. H.O. mit den Defektindizes m, n. Zu einem max. H.O. kann R stets fortgesetzt werden. Wenn $m \neq n$ ist, so gibt es unter seinen Forts. keinen hypermax.; wenn $m=n$ ist, so ist, falls diese endlich sind, jede max. Forts. hypermax., wenn sie $=\infty$ sind, so gibt es unter den max. Forts. sowohl solche, die hypermax. sind, als auch solche, die es nicht sind.*

Beweis. Nach Satz 33 bedeutet Max.: mindestens ein Defektindex $=0$, Hypermax.: beide $=0$, Max. aber nicht Hypermax.: genau einer $=0$. Wenn man hierauf Satz 34 anwendet, gewinnt man genau die Behauptungen unseres Satzes.

Zusatz. *Wenn R nicht schon max. ist, so ist die max. Forts. (und, wenn es solche gibt, die hypermax. sowie die max. aber nicht hypermax. Forts.) auf Kontinuum viele verschiedene Arten möglich.*

Beweis. Unsere Forts.-Methoden kamen (nach Satz 27) auf das Erweitern der $\mathfrak{H}-\mathfrak{E}$, $\mathfrak{H}-\mathfrak{F}$ um zwei abg. lin. M. $\mathfrak{M}, \mathfrak{N}$ heraus, die bei nicht max. R >0 Dimensionen hatten. Aber dabei spielte noch eine längentreue Abbildung U von $\mathfrak{M}$ auf $\mathfrak{N}$ eine Rolle, und diese kann auf Kontinuum viele Arten gewählt werden: zu jedem $\mathfrak{N}$ aufspannenden norm. orth. System $\psi_1, \ldots, \psi_p$ gehörte (nach Satz 26) ein U, und derer gibt es Kontinuum viele (weil z. B. ψ_1 durch jedes $a\psi_1$, $|a|=1$, ersetzt werden kann)[47].

[47]) Mehr als Kontinuum viele Forts. kann es nicht geben, da es nur Kontinuum viele abg. Operatoren in $\mathfrak{H}$ gibt — wie man aus der Separabilität von $\mathfrak{H}$ leicht folgert.

Wenn R nur H. O. ist, so sind seine max. und hypermax. Forts., wie wir wissen, dieselben wie beim abg. lin. H. O. $\tilde{R}$, auf welchen die soeben erzielten Resultate anwendbar sind. Hat insbesondere R nur eine max. Forts., so muß schon $\tilde{R}$ max. sein, ebenso bei den hypermax. Forts.

Wir wenden uns nunmehr der näheren Untersuchung der hypermax. H. O. zu (Kap. IX), und dann der der max., aber nicht hypermax. H. O. (Kap. X).

IX. Die hypermaximalen Operatoren und die Eigenwertdarstellung.

Wenn R hypermaximal ist, so ist $\mathfrak{E} = \mathfrak{F} = \mathfrak{H}$, d. h. die Cayleysche Transformierte U ist unitär: sie ist überall sinnvoll und hat eine (ebenfalls unitäre) überall sinnvolle Inverse U^{-1}. Die hypermax. H. O. R entsprechen als ein-eindeutig den unitären Operatoren U, die die Bedingung von Satz 24 erfüllen, d. h. für die die Menge der $\varphi - U\varphi$ überall dicht ist.

Schon beim Beweise von Satz 24 konstatierten wir: wenn die $\varphi - U\varphi$ überall dicht liegen, so folgt aus $U\varphi = \varphi$ $\varphi = 0$. Es gilt aber hier auch die Umkehrung. Wenn nämlich die $\varphi - U\varphi$ nicht überall dicht liegen, so spannen sie eine abg. lin. M. $\mathfrak{M} \neq \mathfrak{H}$ auf (sie selber bilden schon eine lin. Mann.!) — dann ist $\mathfrak{H} - \mathfrak{M} \neq (0)$. Sei also $f \neq 0$ ein Element von $\mathfrak{H} - \mathfrak{M}$. Alle $\varphi - U\varphi$ gehören zu $\mathfrak{M}$, sind also zu f orth.:

$$(f, \varphi - U\varphi) = (f, \varphi) - (f, U\varphi) = (Uf, U\varphi) - (f, U\varphi) = (Uf - f, U\varphi)$$

muß für alle φ verschwinden; daher ist $Uf - f = 0$, $Uf = f$.

Die in Frage kommenden unitären U können also auch so charakterisiert werden: für sie hat $Uf = f$ keine andere Lösung als die 0 (1 ist nicht „Punkteigenwert“!). Wir beherrschen somit die hypermax. H. O. R, wenn wir diese U kennen. Dies ist aber ein Problem für beschränkte Operatoren (U ist längentreu, also stetig), das auch mit den Methoden der Hilbertschen Theorie der beschränkten Formen (mit einem, auch jenem Ideenkreise angehörenden, Zusatz) bewältigt werden kann. Wir führen die notwendigen Überlegungen (in teilweiser Anlehnung an die einfache Methode von F. Riesz) im Anhang II durch und geben hier nur das Resultat an.

Definition 17. $a < x < b$ sei ein offenes Intervoll (im folgenden wird a, b einmal 0, 1 und einmal $-\infty, \infty$ sein). Unter einer Zerlegung der Einheit (kurz: Z. d. E.) verstehen wir eine Schar von P. O. $E(\lambda)$ (λ läuft durch ganz a, b) mit den folgenden Eigenschaften:

α) Wenn $\lambda \leqq \mu$ ist, so ist $E(\lambda)$ Teil von $E(\mu)$.

β) Wenn $\lambda \geqq \lambda_0$, $\lambda \to \lambda_0$, so gilt für jedes f $E(\lambda)f \to E(\lambda_0)f$.

γ) Wenn $\lambda \to a$ oder $\to b$, so ist $E(\lambda)f \to 0$ bzw. $\to f$.

Wenn nun a, b das Intervall $0, 1$ ist und $E(\varrho)$ eine Z. d. E., so gibt es zu jedem f genau ein f^*, so daß für alle g

$$(f^*, g) = \int_0^1 e^{2\pi i \varrho} d(E(\varrho) f, g) \text{ [48]}$$

gilt. Der durch $Uf = f^*$ definierte Operator U ist unitär, und $Uf = f$ hat keine Lösung außer 0; und umgekehrt wird jedes U mit diesen Eigenschaften durch genau eine Z. d. E. $E(\lambda)$ auf die soeben skizzierte Weise erzeugt (vgl., wie gesagt, Anhang II).

Wir beweisen nun:

Satz 36. *a, b sei das Intervall $-\infty, \infty$ und $F(\lambda)$ eine Z. d. E. Wenn das Integral*

$$\int_{-\infty}^{\infty} \lambda^2 d|F(\lambda) f|^2$$

endlich ist[49]*, so gibt es genau ein f^*, so daß für alle g*

$$(f^*, g) = \int_{-\infty}^{\infty} \lambda \, d(F(\lambda) f, g)$$

(das Integral rechts ist absolut konvergent) gilt. Wir definieren einen Operator R für diese f, und zwar durch $Rf = f^*$.

Dieses R ist ein hypermax. H. O., und jeder hypermax. H. O. R kann mit Hilfe genau einer Z. d. E. auf diese Weise erzeugt werden.

Beweis. Den Z. d. E. $E(\varrho)$ fürs Intervall $0, 1$ werden die Z. d. E. $F(\lambda)$ fürs Intervall $-\infty, \infty$ durch die Zuordnung

$$E(\varrho) \Rightarrow F(\lambda) = E\left(-\frac{1}{\pi} \operatorname{arc\,ctg} \lambda\right), \qquad F(\lambda) \Rightarrow E(\varrho) = F(-\operatorname{ctg} \pi \varrho) \text{ [50]}$$

ein-eindeutig zugeordnet. Es genügt nun zu zeigen: Wenn $E(\lambda)$ die Z. d. E. eines unitären Operators U ist (vgl. die Ausführungen vor unserem Satze), so erzeugt $F(\lambda)$ nach der Anweisung des Satzes wirklich einen Operator R, und dieser ist die Cayleysche Transformierte von U. Nach dem, was wir über das Verhältnis der Cayleyschen Transformierten wissen, und über hypermax. H. O. sowie unitäre Operatoren und Z. d. E. vor dem Satze

---

[48]) Stieljessches Integral, es ist absolut konvergent. Wir schreiben ϱ statt λ.

[49]) Mit λ wächst $F(\lambda)$, also auch $|F(\lambda) f|^2$, außerdem ist λ^2 nie negativ, daher ist dieses Integral seiner Natur nach entweder konvergent (endlich) oder eigentlich divergent ($+\infty$).

[50]) Durch $\varrho = -\frac{1}{\pi} \operatorname{arc\,ctg} \lambda$, $\lambda = -\operatorname{ctg} \pi \varrho$ werden die Intervalle $0 < \varrho < 1$ und $-\infty < \lambda < \infty$ ein-eindeutig aufeinander bezogen. (Beim arc ctg ist immer der Wert $> -\pi$, < 0 zu nehmen.)

formuliert haben, folgt hieraus tatsächlich die Richtigkeit aller Behauptungen.

Sei also $F(\lambda)$ eine Z. d. E., $E(\varrho) = F(-\operatorname{ctg} \pi \varrho)$, U der dazugehörige unitäre Operator. Zunächst zeigen wir: Wenn $\int\limits_{-\infty}^{\infty} \lambda^2 d|F(\lambda)f|^2 = C^2$ endlich ist, so kann das f^* unseres Satzes wirklich erzeugt werden.

Es sei $\lambda < \lambda' < \mu$. Da $E(\lambda)$ Teil von $E(\mu)$ ist, ist $E(\mu) - E(\lambda)$ ein P. O. und deshalb

$$|(\{E(\mu) - E(\lambda)\}f, g)| = |(\{E(\mu) - E(\lambda)\}f, \{E(\mu) - E(\lambda)\}g)|$$
$$\leqq |\{E(\mu) - E(\lambda)\}f| \cdot |\{E(\mu) - E(\lambda)\}g|,$$

und weil allgemein

$$|\{E(\mu) - E(\lambda)\}h|^2 = (h, \{E(\mu) - E(\lambda)\}h) = (h, E(\mu)h) - (h, E(\lambda)h)$$
$$= |E(\mu)h|^2 - |E(\lambda)h|^2$$

gilt,

$$|(\{E(\mu) - E(\lambda)\}f, g)| \leqq \sqrt{|E(\mu)f|^2 - |E(\lambda)f|^2}\,\sqrt{|E(\mu)g|^2 - |E(\lambda)g|^2},$$
$$|\lambda'\{(E(\mu)f, g) - (E(\lambda)f, g\})$$
$$\leqq \sqrt{\lambda'^2\{|E(\mu)f|^2 - |E(\lambda)f|^2\}}\,\sqrt{|E(\mu)g|^2 - |E(\lambda)g|^2}.$$

Da nun die beiden Integrale $\int\limits_{-\infty}^{\infty} \lambda^2 d|E(\lambda)f|^2 = C^2$ und $\int\limits_{-\infty}^{\infty} d|E(\lambda)g|^2 = |E(\lambda)g|^2 \Big|_{\lambda=-\infty}^{\lambda=\infty} = |g|^2$ absolut konvergieren, hat die bekannte Ungleichung

$$\sqrt{a_1 b_1} + \dots + \sqrt{a_p b_p} \leqq \sqrt{(a_1 + \dots + a_p)(b_1 + \dots + b_p)}$$
$$(\text{alle } a_1, \dots, a_p,\ b_1, \dots, b_p \geqq 0)$$

erstens die Konsequenz, daß auch $\int\limits_{-\infty}^{\infty} \lambda\, d(E(\lambda)f, g)$ absolut konvergiert, und zweitens daß

$$\Big|\int\limits_{-\infty}^{\infty} \lambda\, d(E(\lambda), g)\Big| \leqq \sqrt{\int\limits_{-\infty}^{\infty} \lambda^2 d|E(\lambda)f|^2 \cdot \int\limits_{-\infty}^{\infty} d|E(\lambda)g|^2} = C \cdot |g|$$

ist[51]). Wir nennen dieses Integral für einen Augenblick $L(g)$ (f fest, g variabel), dann hat $L(g)$ die Eigenschaften:

$$L(a_1 g_1 + \dots + a_n g_n) = \bar{a}_1 L(g_1) + \dots + \bar{a}_n L(g_n), \quad |L(g)| \leqq C \cdot |g|$$

(es ist linear und stetig). Nach einem Satze von F. Riesz gibt es darum

---

[51]) Es genügt, sich die Definition des Stieltjesschen Integrals zu vergegenwärtigen. Die genannte Ungleichung ist nur eine veränderte Schreibweise der Schwarzschen.

genau ein f^*, so daß stets $(f^*, g) = L(g)$ ist[52]) — d. i. gerade unsere auf f^* bezügliche Behauptung.

R ist also wirklich herstellbar. Sei R^* die Cayleysche Transformierte von U, wir müssen noch $R = R^*$ beweisen. Es genügt aber zu zeigen, daß R Forts. von R^* ist. Denn wenn Rf, Rg Sinn haben, so geht (nach Definition in Satz 36) beim Vertauschen von f, g (Rf, g) in seine komplex-konjugierte über — also ist $(Rf, g) = (f, Rg)$. Daher ist R ein H. O (sein Definitionsbereich enthält nach Annahme den von R^*, ist also überall dicht) und Forts. des (wegen der Unitarität von U) max. R^*, also $R = R^*$.

Sei also $R^* f$ sinnvoll, wir müssen beweisen: Rf ist sinnvoll und $= R^* f$; oder was dasselbe ist: $\int\limits_{-\infty}^{\infty} \lambda^2 d\,|F(\lambda) f|^2$ endlich und für alle g $(R^* f, g) = (Rf, g) = \int\limits_{-\infty}^{\infty} \lambda\, d(F(\lambda) f, g)$. Daß Rf sinnvoll ist, bedeutet übrigens $f = \varphi - U\varphi$, dieses sei unser Ausgangspunkt.

$\int\limits_{-\infty}^{\infty} \lambda^2 d\,|F(\lambda) f|^2$, $\int\limits_{-\infty}^{\infty} \lambda\, d(F(\lambda) f, g)$ können wir, nach der Transformation $\lambda = -\operatorname{ctg} \pi\varrho$, auch so schreiben: $\int\limits_0^1 \operatorname{ctg}^2 \pi\varrho\, d\,|E(\varrho) f|^2$, $\int\limits_0^1 -\operatorname{ctg} \pi\varrho\, d(E(\varrho) f, g)$. Um hier $f = \varphi - U\varphi$ einsetzen zu können, müssen wir zunächst einige Integrale ausrechnen.

$$(\varphi \pm U\varphi, g) = \int\limits_0^1 d(E(\varrho)\varphi, g) \pm \int\limits_0^1 e^{2\pi i \cdot \varrho} d(E(\varrho)\varphi, g)$$
$$= \int\limits_0^1 (1 \pm e^{2\pi i \cdot \varrho})\, d(E(\varrho)\varphi, g),$$

---

[52]) Man beweist diesen Satz so:

Da alle (f^*, g) vorgeschrieben sind, kann es nicht zwei verschiedene f^* geben, es genügt also die Existenz von einem zu beweisen. Sei $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System, $L(\varphi_n) = a_n$ $(n = 1, 2, \ldots)$. Es ist $g = \sum\limits_{n=1}^{\infty} (g, \varphi_n) \cdot \varphi_n$ (Satz 7, β)), also wegen der Linearität und Stetigkeit von $L(g)$ $L(g) = \sum\limits_{n=1}^{\infty} \overline{(g, \varphi_n)} \cdot a_n$ (die Reihe muß konvergieren). Wäre nun $\sum\limits_{n=1}^{\infty} |a_n|^2$ endlich, so würde $f^* = \sum\limits_{n=1}^{\infty} a_n \varphi_n$ konvergieren (Satz 5), $(f^*, \varphi_n) = a_n$, und daraus folgt $L(g) = \sum\limits_{n=1}^{\infty} (f^*, \varphi_n) \cdot \overline{(g, \varphi_n)} = (f^*, g)$ (Satz 7 γ)): d. i. die Behauptung.

Nun ist $|L(g)| \leqq C \cdot |g|$, also für $g = x_1 \varphi_1 + \ldots + x_m \varphi_m$

$$\left| \sum_{n=1}^{m} a_n \bar{x}_n \right| \leqq C \cdot \sqrt{\sum_{n=1}^{m} |x_n|^2},$$

und wenn wir $x_n = a_n$ $(n = 1, \ldots, m)$ setzen: $\sum\limits_{n=1}^{m} |a_n|^2 \leqq C^2$. Daher ist $\sum\limits_{n=1}^{\infty} |a_n|^2$ endlich (und zwar $\leqq C^2$).

$$
\begin{aligned}
(\varphi - U\varphi, E(\varrho)g) &= \int_0^1 (1 - e^{2\pi i\cdot\varrho'}\, d(E(\varrho'))\varphi, E(\varrho)g) \\
&= \int_0^1 (1 - e^{2\pi i\cdot\varrho'})\, d(E(\varrho)E(\varrho')\varphi, g) \\
&= \int_0^1 (1 - e^{2\pi i\cdot\varrho'})\, d(E(\mathrm{Min}(\varrho, \varrho'))\varphi, g) \text{ [53]} \\
&= \int_0^\varrho (1 - e^{2\pi i\cdot\varrho'})\, d(E(\varrho')\varphi, g),
\end{aligned}
$$

$$
\begin{aligned}
|E(\varrho)(\varphi - U\varphi)|^2 &= (\varphi - U\varphi, E(\varrho)(\varphi - U\varphi)) \\
&= \int_0^\varrho (1 - e^{2\pi i\cdot\varrho'})\, d(E(\varrho')\varphi, \varphi - U\varphi) \text{ [54]} \\
&= \int_0^\varrho (1 - e^{2\pi i\cdot\varrho'})\, d\left[\int_0^{\varrho'} (1 - e^{-2\pi i\cdot\varrho''})\, d(E(\varrho'')\varphi, \varphi)\right] \text{ [55]} \\
&= \int_0^\varrho (1 - e^{2\pi i\cdot\varrho'})(1 - e^{-2\pi i\cdot\varrho'})\, d(E(\varrho')\varphi, \varphi) \\
&= \int_0^\varrho 4\sin^2\pi\varrho'\, d\,|E(\varrho')\varphi|^2,
\end{aligned}
$$

$$
\begin{aligned}
\int_0^1 \mathrm{ctg}^2\,\pi\varrho\, d\,|E(\varrho)(\varphi - U\varphi)|^2 &= \int_0^1 \mathrm{ctg}^2\,\pi\varrho\, d\left[\int_0^\varrho 4\sin^2\pi\varrho'\, d\,|E(\varrho')\varphi|^2\right] \text{ [55]} \\
= \int_0^1 \mathrm{ctg}^2\,\pi\varrho\cdot 4\sin^2\pi\varrho\, d\,|E(\varrho)\varphi|^2 &= \int_0^1 4\cos^2\pi\varrho\, d\,|E(\varrho)\varphi|^2,
\end{aligned}
$$

und dies ist endlich, weil es durch das endliche $\int_0^1 4\, d\,|E(\varrho)\varphi|^2 = 4\,|\varphi|^2$ majorisiert wird. Weiter ist

$$
\int_0^1 -\mathrm{ctg}\,\pi\varrho\, d(E(\varrho)(\varphi - U\varphi), g) = \int_0^1 -\mathrm{ctg}\,\pi\varrho\, d\left[\int_0^\varrho (1 - e^{2\pi i\cdot\varrho'})\, d(E(\varrho')\varphi, g)\right] \text{ [55]}
$$

$$
\begin{aligned}
&= \int_0^1 -\mathrm{ctg}\,\pi\varrho\cdot(1 - e^{2\pi i\cdot\varrho})\, d(E(\varrho)\varphi, g) = \int_0^1 i(1 + e^{2\pi i\cdot\varrho})\, d(E(\varrho)\varphi, g) \\
&= (i(\varphi + U\varphi), g) = (R^*(\varphi - U\varphi), g).
\end{aligned}
$$

Damit ist alles Erforderliche bewiesen. –

---

[53] Für $\varrho' > \varrho$ ist ja $(E(\mathrm{Min}(\varrho, \varrho'))\varphi, g)$ konstant, also das Integral 0!

[54] Man nehme mit $g = \varphi$ das komplex-Konjugierte der letzten Formel.

[55] Nach einer allgemeinen Formel für Stieltjessche Integrale ist

$$
\int_A^B p(\varrho)\cdot d\left[\int_A^\varrho q(\varrho')\cdot dr(\varrho')\right] = \int_A^B p(\varrho)\cdot q(\varrho)\cdot dr(\varrho).
$$

Wir haben für hypermax. H. O. die der Hilbertschen Spektraldarstellung oder Eigenwertdarstellung (bei beschränkten Formen, vgl. das Zitat von Anm. [7])) analoge Normalform gewonnen, dabei aber für diese bereits alle Eigenwertdarstellungen verbraucht, so daß es für max. und nicht hypermax. H. O. gewiß keine solchen geben kann. (Allerdings haben wir noch kein Beispiel eines derartigen H. O. konstruiert.) Darum sehen wir von der kaum neue Gesichtspunkte bietenden weiteren Untersuchung der hypermax. H. O. ab und wenden uns den max., aber nicht hypermax. H. O. zu.

X. Maximale, aber nicht hypermaximale Operatoren.

Ein solcher H. O. R (m, n seine Defektindizes) ist durch $m = 0$, $n > 0$ oder $m > 0$, $n = 0$ charakterisiert; da aber beim Ersetzen von R durch $-R$ m, n vertauscht werden, genügt es $m = 0$, $n > 0$ zu betrachten. Wir bilden $U, \mathfrak{E}, \mathfrak{F}$, es ist $\mathfrak{E} = \mathfrak{H}$ und $\mathfrak{G}_0 = \mathfrak{H} - \mathfrak{F}$ ist n-dimensional, U bildet $\mathfrak{H}$ längentreu auf $\mathfrak{F}$ ab.

Das durch U^ν $(\nu = 0, 1, 2, \ldots)$ vermittelte Bild von $\mathfrak{G}_0$ sei $\mathfrak{G}_\nu$ (auch eine abg. lin. M.). Für $\nu > 0$ ist $\mathfrak{G}_\nu$ Teil von $\mathfrak{F}$, also orth. zu $\mathfrak{G}_0$; wenn wir hierauf die Abbildung U^k $(k = 0, 1, 2, \ldots)$ anwenden, so werden die Bilder $\mathfrak{G}_{k+\nu}, \mathfrak{G}_k$ zueinander orth. (weil (f, g) dabei invariant ist, vgl. Satz 10). Also sind allgemein $\mathfrak{G}_k, \mathfrak{G}_l$ $(k, l = 0, 1, 2, \ldots)$ für $k \neq l$ orth. Die $\mathfrak{G}_0, \mathfrak{G}_1, \mathfrak{G}_2, \ldots$ mögen die abg. lin. M. $\mathfrak{M}$ aufspannen, wir setzen $\mathfrak{H} - \mathfrak{M} = \mathfrak{N}$.

Daß U $\mathfrak{G}_\nu$ auf $\mathfrak{G}_{\nu+1}$ abbildet (natürlich ein-eindeutig, längentreu) wissen wir schon, wir wollen nun zeigen, daß $\mathfrak{N}$ (ebenso) auf sich selbst abgebildet wird. $\mathfrak{N}$ ist, wie man leicht einsieht, die Menge aller zu allen $\mathfrak{G}_\nu$ orth. Elemente von $\mathfrak{H}$, sein Bild also (wegen der Invarianz von (f, g)) die Menge aller zu allen Bildern der $\mathfrak{G}_\nu$ orth. Elemente des Bildes von $\mathfrak{H}$: dies sind die zu allen $\mathfrak{G}_{\nu+1}$ orth. Elemente von $\mathfrak{F} = \mathfrak{H} - \mathfrak{G}_0$, also die zu $\mathfrak{G}_0$ und allen $\mathfrak{G}_{\nu+1}$ orth. Elemente von $\mathfrak{H}$. Also wirklich wieder $\mathfrak{N}$.

Sei nun $\varphi_1^0, \ldots, \varphi_p^0$ ein die abg. lin. M. $\mathfrak{G}_0$ aufspannendes norm. orth. System (eventuell $p = \infty$!), und $U^\nu \varphi_q^0 = \varphi_q^\nu$ $(\nu = 0, 1, 2, \ldots,\ q = 1, \ldots, p)$, dann bilden die $\varphi_1^\nu, \ldots, \varphi_p^\nu$ auch ein norm. orth. System und spannen $\mathfrak{G}_\nu$ auf. Für $k \neq l$ liegt $\varphi_{q'}^k$ in $\mathfrak{G}_k$, $\varphi_{q''}^l$ in $\mathfrak{G}_l$, also sind beide orth., d. h. alle φ_q^ν $(\nu = 0, 1, 2, \ldots, q = 1, \ldots, p)$ bilden auch ein norm. orth. System — und zwar spannen sie dasselbe auf, wie die $\mathfrak{G}_\nu$: $\mathfrak{M}$.

$\varphi_q^0, \varphi_q^1, \varphi_q^2, \ldots$ mögen $\mathfrak{M}_q$ aufspannen, $\varphi_q^1, \varphi_q^2, \ldots$, $\overline{\mathfrak{M}}_q$ (alles als abg. lin. M., $q = 1, \ldots, p$), dann spannen $\mathfrak{M}_1, \ldots, \mathfrak{M}_p$ wieder $\mathfrak{M}$ auf, und $\overline{\mathfrak{M}}_q$ ist Teil von $\mathfrak{M}_p$. Da $U\varphi_q^\nu = \varphi_q^{\nu+1}$ ist, bildet U $\mathfrak{M}_q$ auf $\overline{\mathfrak{M}}_q$ ab. Zusammenfassend haben wir also:

$\mathfrak{M}_1, \ldots, \mathfrak{M}_p$, $\mathfrak{N}$ spannen $\mathfrak{H}$ auf (wie $\mathfrak{M}, \mathfrak{N}$), U bildet sie bzw. auf $\overline{\mathfrak{M}}_1, \ldots, \overline{\mathfrak{M}}_p$, $\mathfrak{N}$ ab, d. h. auf Teilmengen von sich selbst. Ferner sind die $\mathfrak{M}_1, \ldots, \mathfrak{M}_p$ (als Teile von $\mathfrak{M}$) zu $\mathfrak{N}$ orth., und auch zueinander (weil es die $\varphi_{q'}^k, \varphi_{q''}^l$ mit $q' \neq q''$ sind). Hieraus folgt, daß jedes $\mathfrak{M}_1, \ldots, \mathfrak{M}_p$, $\mathfrak{N}$ das U reduziert, also (Satz 25) auch die Cayleysche Transformierte R. —

Wir betrachten die in $\mathfrak{M}_1, \ldots, \mathfrak{M}_p$, $\mathfrak{N}$ liegenden Teile von R und U: es sind bzw. $R_1, \ldots, R_p, S$ und bzw. $U_1, \ldots, U_p, V$; R ist mit Hilfe von $R_1, \ldots, R_p, S$ nach Satz 21 zu charakterisieren. Nun sind die abg. lin. M. $\mathfrak{M}_1, \ldots, \mathfrak{M}_p$ alle unendlichvieldimensional ($\mathfrak{M}_q$ wird ja vom norm. orth. System $\varphi_q^0, \varphi_q^1, \varphi_q^2, \ldots$ aufgespannt), also lauter Hilbertsche Räume (d. h. sie genügen, bei unveränderter Definition von $af, f+g, (f, g)$ den Bedingungen A bis E von Kap. I [56])), daher dürfen wir in ihnen alle für $\mathfrak{H}$ ausgebildeten Begriffe ohne weiteres verwenden. Also ist R_q in $\mathfrak{M}_q$ ein H. O. (insbesondere ist sein Definitionsbereich in $\mathfrak{M}_q$ überall dicht, denn er besteht aus den Projektionen des in $\mathfrak{H}$ überall dichten Definitionsbereiches von R), U_q ist längentreu und die Cayleysche Transformierte von R_q. In $\mathfrak{N}$ sind drei Fälle möglich: erstens kann es ∞-dimensional, also ein Hilbertscher Raum sein, dann ist wieder S ein H. O., V unitär (es bildet ja $\mathfrak{N}$ auf $\mathfrak{N}$ ab) und seine Cayleysche Transformierte — also S hypermax.; zweitens kann $\mathfrak{N}$ N $(=1, 2, \ldots)$-dimensional, also ein N-dimensionaler (komplexer) Euklidischer Raum sein (vgl. Anm. [56])), dann ist S ein H. O. in ihm — d. h. ein gewöhnlicher, endlichvieldimensionaler Hermitescher Operator; drittens kann $\mathfrak{N}$ 0-dimensional sein, d. h. $=(0)$ — dann brauchen wir es gar nicht zu berücksichtigen.

Bei $\mathfrak{N}$ und S geschieht also sicher nichts Neues, wir wenden uns darum den $\mathfrak{M}_q, R_q$ und U_q $(q=1, \ldots, p)$ zu. Diese benehmen sich alle gleich: R_q ist Cayleysche Transformierte von U_q, und es gibt ein in $\mathfrak{M}_q$ vollst. (d. h. $\mathfrak{M}_q$ aufspannendes) norm. orth. System $\varphi_q^0, \varphi_q^1, \varphi_q^2, \ldots$, so daß $U_q \varphi_q^\nu = U \varphi_q^\nu = \varphi_q^{\nu+1}$ ist, also $U_q\left(\sum_{\nu=0}^{\infty} x_\nu \varphi_q^\nu\right) = \sum_{\nu=0}^{\infty} x_\nu \varphi_q^{\nu+1}$ $\left(\sum_{\nu=0}^{\infty} |x_\nu|^2 \text{ endlich}\right)$; hierdurch ist aber U_q und somit R_q völlig festgelegt (wenn man $\varphi_q^0, \varphi_q^1, \varphi_q^2, \ldots$ als bekannt ansieht). Indem wir die Isomorphie von $\mathfrak{M}_q$ mit $\mathfrak{H}$ ausnützen, können wir auch sagen, daß R_q mit dem folgenden Operator $\overline{R}$ in $\mathfrak{H}$ isomorph sein muß: $\overline{R}$ ist die Cayleysche Transformierte von $\overline{U}$, und $\overline{U}$ ist

[56]) A bis C und E genügt offenbar jede abg. lin. M., D aber ist der Unendlichviel-dimensionalität gleichwertig. Eine endlichviel-dimensionale abg. lin. M. besteht, falls sie 0 Dimensionen hat, aus der 0 allein; wenn sie $N = 1, 2, \ldots$ Dimensionen hat, so wird sie durch ein norm. orth. System $\varphi_1, \ldots, \varphi_N$ aufgespannt, sie besteht also aus den $x_1\varphi_1 + \ldots + x_N\varphi_N$, d. h. sie ist der N-dimensionale (komplexe) Euklidische Raum aller $x_1, \ldots, x_N$.

(wenn $\psi_0, \psi_1, \psi_2, \ldots$ ein festes vollst. norm. orth. System in $\mathfrak{H}$ ist) durch $\overline{U}\left(\sum_{\nu=0}^{\infty} x_\nu \psi_\nu\right) = \sum_{\nu=0}^{\infty} x_\nu \psi_{\nu+1}$ $\left(\sum_{\nu=0}^{\infty} |x_\nu|^2 \text{ endlich}\right)$ definiert. —

Dabei wäre aber zuerst noch zu zeigen, daß $\overline{R}$ wirklich existiert, d. h. daß das offenbar längentreue $\overline{U}$ der Bedingung von Satz 24 genügt. Wenn das der Fall ist, so ist $\overline{R}$, da $\overline{U}$ in ganz $\mathfrak{H}$ definiert (also $\mathfrak{E} = \mathfrak{H}$, $m = 0$) ist, ein max. H. O. Der Wertevorrat von $\overline{U}$, $\mathfrak{F}$ ist die Menge aller $\sum_{\nu=0}^{\infty} x_\nu \psi_{\nu+1}$, d. h. die von den $\psi_1, \psi_2, \ldots$ aufgespannte abg. lin. M. — also ist $\mathfrak{H} - \mathfrak{F}$ die Menge aller $a\psi_1$, also 1-dimensional. Folglich hat $\overline{R}$ die Defektindizes 0, 1: es ist das erste Beispiel eines max., aber nicht hypermax. H. O.

Zunächst ist aber noch seine Existenz zu beweisen, d. h. die überall Dichtigkeit der Menge aller $\varphi - \overline{U}\varphi$. Da sie eine lin. M. ist, genügt es zu zeigen, daß alle Elemente eines vollst. norm. orth. Systems Häufungspunkte von ihr sind — etwa alle ψ_ν. Nun ist

$$\left(\psi_\nu + \frac{l-1}{l}\psi_{\nu+1} + \cdots + \frac{1}{l}\psi_{\nu+l-1}\right) - \overline{U}\left(\psi_\nu + \frac{l-1}{l}\psi_{\nu+1} + \cdots + \frac{1}{l}\psi_{\nu+l-1}\right)$$
$$= \left(\psi_\nu + \frac{l-1}{l}\psi_{\nu+1} + \cdots + \frac{1}{l}\psi_{\nu+l-1}\right) - \left(\psi_{\nu+1} + \frac{l-1}{l}\psi_{\nu+2} + \cdots + \frac{1}{l}\psi_{\nu+l}\right)$$
$$= \psi_\nu - \frac{1}{l}(\psi_{\nu+1} + \cdots + \psi_{\nu+l}),$$

und dies unterscheidet sich von ψ_ν um $\frac{1}{l}(\psi_{\nu+1} + \cdots + \psi_{\nu+l})$, wobei

$$\left|\frac{1}{l}(\psi_{\nu+1} + \cdots + \psi_{\nu+l})\right|^2 = \frac{1}{l}, \quad \left|\frac{1}{l}(\psi_{\nu+1} + \cdots + \psi_{\nu+l})\right| = \frac{1}{\sqrt{l}}$$

ist, also (bei geeignetem l) beliebig klein. Damit ist alles bewiesen.

Wir formulieren:

Satz 37. *Sei $\psi_0, \psi_1, \psi_2, \ldots$ ein vollst. norm. orth. System. Es gibt dann genau einen H. O. $\overline{R}$, dessen Cayleysche Transformierte $\overline{U}$ durch*

$$\overline{U}\left(\sum_{\nu=0}^{\infty} x_\nu \psi_\nu\right) = \sum_{\nu=0}^{\infty} x_\nu \psi_{\nu+1} \qquad \left(\sum_{\nu=0}^{\infty} |x_\nu|^2 \textit{ endlich}\right)$$

definiert ist. $\overline{R}$ hat die Defektindizes 0, 1, *es ist also max., aber nicht hypermax.*

Beweis. Wurde soeben erbracht. —

Wir werden im folgenden noch einige wichtige allgemeine Eigenschaften von $\overline{R}$ herleiten, seiner expliziten Darstellung ist der Teil VI der „Bemerkungen“ gewidmet. Zunächst greifen wir auf unsere früheren Entwickelungen zurück und zeigen:

Satz 38. *R sei ein max., aber nicht hypermax. H. O., also seine Defektindizes* $0, n$ $(n = 1, 2, \ldots, \infty)$ *oder* $m, 0$ $(m = 1, 2, \ldots, \infty)$. *Es sei*

$p = n$ bzw. $= m$, wir können dann $\mathfrak{H}$ in $p+1$ abg. lin. M. $\mathfrak{M}_1, \ldots, \mathfrak{M}_p, \mathfrak{N}$ zerlegen, die paarweise orth. sind und zusammen die abg. lin. M. $\mathfrak{H}$ aufspannen (für $p = \infty$ bricht die Folge $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ nicht ab), so daß

α) jedes $\mathfrak{M}_1, \ldots, \mathfrak{M}_p, \mathfrak{N}$ R reduziert, der in ihnen liegende Teil von R sei bzw. $R_1, \ldots, R_p, S$;

β) jedes $\mathfrak{M}_q$ $(q = 1, 2, \ldots, p)$ ist unendlichvieldimensional und ein Hilbertscher Raum, d. h. $\mathfrak{H}$ isomorph; und in ihnen sind alle R_q dem $\bar{R}$ isomorph bzw. alle R_q dem $-\bar{R}$ isomorph ($\bar{R}$ nach Satz 37).

γ) Für $\mathfrak{N}$ bestehen drei Möglichkeiten: erstens kann es unendlichvieldimensional und ein Hilbertscher Raum, d. h. $\mathfrak{H}$ isomorph sein, dann ist S ein hypermax. H. O.; zweitens kann es N $(=1, 2, \ldots)$-dimensional, also ein N-dimensionaler (komplexer) Euklidischer Raum, sein, dann ist S ein gewöhnlicher endlichvieldimensionaler Hermitescher Operator; drittens kann es 0-dimensional, d. h. $=(0)$ sein, dann kann man es, zusammen mit S, fortlassen.

Diese $R_1, \ldots, R_p, S$ bestimmen R nach Satz 21.

Beweis. Wurde für $m = 0$, $n > 0$ schon im ersten Teile dieses Kapitels erbracht (man beachte noch Satz 37), für $m > 0$, $n = 0$ gewinnt man ihn durch Ersetzen von R durch $-R$ (also Vertauschen von m, n). —

In endlichvieldimensionalen Räumen kann jeder Hermitesche Operator auf die Diagonalform gebracht werden, solange er >1 Dimensionen hat ist er also reduzibel. In $\mathfrak{H}$ sind ebenfalls die hypermax. H. O. reduzibel, wovon man sich wie folgt überzeugt.

Sei R hypermax., $F(\lambda)$ seine Z. d. E. nach Satz 36. Nehmen wir zuerst an, es gäbe ein λ, so daß $F(\lambda) \neq 0, 1$ ist, also die abg. lin. M. $\mathfrak{M}$ mit $P_{\mathfrak{M}} = F(\lambda) \neq (0), \mathfrak{H}$. Dieses $\mathfrak{M}$ reduziert nun R, denn es ist erstens

$$\int_{-\infty}^{\infty} \lambda'^2 \, d\,|E(\lambda')E(\lambda)f|^2 = \int_{-\infty}^{\infty} \lambda'^2 \, d\,|E(\mathrm{Min}(\lambda', \lambda))f|^2 \quad (\text{vgl. Anm. } ^{53}))$$

$$= \int_{-\infty}^{\lambda} \lambda'^2 \, d\,|E(\lambda')f|^2,$$

und dies ist endlich, wenn $\int_{-\infty}^{\infty} \lambda' \, d\,|E(\lambda')f|^2$ es ist, und zweitens gilt dann

$$(E(\lambda)Rf, g) = (Rf, E(\lambda)g) = \int_{-\infty}^{\infty} \lambda' \, d(E(\lambda')f, E(\lambda)g)$$

$$= \int_{-\infty}^{\infty} \lambda' \, d(E(\lambda)E(\lambda')f, g) = \int_{-\infty}^{\infty} \lambda' \, d(E(\lambda')E(\lambda)f, g) = (RE(\lambda)f, g)$$

für alle g, d. h. es ist $E(\lambda)Rf = RE(\lambda)f$.

Sind aber alle $E(\lambda) = 0$ oder 1, so gibt es (wegen α) in Def. 17) ein λ_0, so daß für $\lambda < \lambda_0$ das erstere und für $\lambda > \lambda_0$ das letztere gilt, λ_0 ist endlich (wegen β)), und für $\lambda = \lambda_0$ gilt das zweite (wegen γ)). Daher ist

$$(Rf, g) = \int_{-\infty}^{\infty} \lambda' d(E(\lambda) f, g) = \lambda_0 (f, g), \qquad Rf = \lambda_0 f,$$

d. h. $\lambda_0 \cdot 1$ eine Forts. von R, also $R = \lambda_0 \cdot 1$. Dieses wird aber offenbar überhaupt durch jede abg. lin. M. reduziert.

Die Reduzibilität ist somit ein Rudiment der Diagonal-Transformierbarkeit und als solches mit der Möglichkeit der Eigenwertdarstellung aufs engste verbunden; Irreduzibilität einer Matrix bedeutet bereits das Auftreten der Ausnahmefälle der Elementarteilertheorie — was durch den Hermiteschen Charakter im Endlichvieldimensionalen, und wie wir sahen auch in $\mathfrak{H}$ im Hypermax., ausgeschlossen wird. Unter diesem Gesichtspunkte ist nun die folgende Eigenschaft von $\overline{R}$ sehr beachtenswert:

Satz 39. *Ein max. H. O. ist dann und nur dann irreduzibel, wenn er (bei geeigneter Wahl des vollst. norm. orth. Systems $\psi_0, \psi_1, \psi_2, \ldots$ aus Satz 37) mit $\overline{R}$ oder mit $-\overline{R}$ zusammenfällt (beides läßt sich nie beim selben H. O. erreichen).*

Beweis. Sei R max. und irreduzibel. Nach dem früher Gesagten ist es dann gewiß nicht hypermax., also tritt Satz 38 in Kraft. Da das dortige $\mathfrak{M}_1$ R reduziert, und $\mathfrak{M}_1 \neq 0$ ist, muß $\mathfrak{M}_1 = \mathfrak{H}$ sein — also $p = 1$ und $\mathfrak{N}$ fehlend. Dann ist $R = R_1$ also nach β) dortselbst eben $\overline{R}$ oder $-\overline{R}$ (bei richtiger Wahl von $\psi_0, \psi_1, \psi_2, \ldots$). Da im einen Falle die Defektindizes von R 0, 1 sind, im anderen 1, 0, ist sicher beides unvereinbar.

Es bleibt noch übrig zu zeigen, daß $\overline{R}$ (und also auch $-\overline{R}$) irreduzibel ist — d. h. daß $\overline{U}$ irreduzibel ist (vgl. Satz 25). Möge also $\overline{U}$ durch $\mathfrak{M}$ reduziert werden, also auch durch $\mathfrak{N} = \mathfrak{H} - \mathfrak{M}$; jede dieser Mengen wird also durch das (überall sinnvolle) $\overline{U}$ auf eine Teilmenge von sich abgebildet: $\overline{\mathfrak{M}}, \overline{\mathfrak{N}}$. Wäre $\overline{\mathfrak{M}} = \mathfrak{M}$, $\overline{\mathfrak{N}} = \mathfrak{N}$, so enthielte der Wertevorrat von U $\mathfrak{M}$ und $\mathfrak{N}$, also auch $\mathfrak{M} \dotplus \mathfrak{N} = \mathfrak{H}$, was nicht der Fall ist. Also ist $\overline{\mathfrak{M}}$ oder $\overline{\mathfrak{N}}$ echtes Teil von $\mathfrak{M}$ oder $\mathfrak{N}$, $\mathfrak{M} - \overline{\mathfrak{M}}$ oder $\mathfrak{N} - \overline{\mathfrak{N}} \neq (0)$. Sei $f \neq 0$ ein Element von $\mathfrak{M} - \overline{\mathfrak{M}}$ bzw. $\mathfrak{N} - \overline{\mathfrak{N}}$. Dann ist es orth. zu $\overline{\mathfrak{M}}$ bzw. $\overline{\mathfrak{N}}$, und als Element von $\mathfrak{M}$ bzw. $\mathfrak{N}$ auch zu $\mathfrak{N}$ bzw. $\mathfrak{M}$ und $\overline{\mathfrak{N}}$ bzw. $\overline{\mathfrak{M}}$: also zum ganzen Wertevorrat von $\overline{U}$, $\overline{\mathfrak{M}} \dotplus \overline{\mathfrak{N}}$. D. h. zu allen $\psi_1, \psi_2, \ldots$, es ist also $= a\psi_0$ (mit $a \neq 0$). Folglich gehört auch ψ_0 zu $\mathfrak{M} - \overline{\mathfrak{M}}$ bzw. $\mathfrak{N} - \overline{\mathfrak{N}}$, also zu $\mathfrak{M}$ bzw. $\mathfrak{N}$. Aber $\mathfrak{M}, \mathfrak{N}$ enthalten auch ihre $\overline{U}$-Bilder, das betreffende von ihnen enthält daher mit ψ_0 auch $\psi_1, \psi_2, \ldots$ — also ganz $\mathfrak{H}$. Darum ist $\mathfrak{M} = \mathfrak{H}$ oder $\mathfrak{H} - \mathfrak{M} = \mathfrak{N} = \mathfrak{H}$, $\mathfrak{M} = (0)$, also alles bewiesen. —

Mit $\bar{R}$ ist auch $a\bar{R} + b1$ (a, b reell, $a \neq 0$) max. und irreduzibel, also (bis aufs Koordinatensystem) $= \bar{R}$ oder $= -\bar{R}$. Für $a > 0$ sind die Defektindizes $0, 1$, also kommt dann $\bar{R}$ in Frage, für $a < 0$ sind sie $1, 0$, also dann $-\bar{R}$.

Zum Schluß sei noch erwähnt, daß die Zerlegung des H. O. R im Satz 38 eine Art Umkehrung zuläßt, die die Konstruktion von abg. lin. H. O. R mit beliebig vorgegebenen Defektindizes m, n ermöglicht. Da $m = n = 0$ durch jeden hypermax. H. O. (also z. B. 1) realisiert wird, können wir $m + n > 0$ annehmen.

Dann gibt es $m + n$ unendlichvieldimensionale abg. lin. M. $\mathfrak{M}_1, \ldots, \mathfrak{M}_m$, $\mathfrak{N}_1, \ldots, \mathfrak{N}_n$, die zu je zweien orth. sind und zusammen $\mathfrak{H}$ aufspannen (als abg. lin. M.)[57]; jede ist dem $\mathfrak{H}$ isomorph, wir können darum in den m ersten $-\bar{R}$ und den n übrigen $\bar{R}$ realisieren: es seien die Operatoren $R_1, \ldots, R_m, S_1, \ldots, S_n$. Man überlegt sich leicht, daß diese, nach der Anweisung von Satz 21 zusammengefaßt, einen abg. lin. H. O. R ergeben, und daß R die Defektindizes m, n hat.

XI. Halbbeschränktheit.

Wir wollen einige Merkmale angeben, die jedem nicht hypermax. (aber max.!) H. O. abgehen müssen. (Die meisten Resultate dieses Satzes gewinnen wir übrigens bald auch auf anderem Wege.)

Satz 40. ***Ein max., aber nicht hypermax. H. O. kann weder nach oben noch nach unten halbbeschränkt (also erst recht nicht beschränkt) sein, er kann weder überall sinnvoll sein noch alle Werte annehmen, und er kann in keinem vollst. norm. orth. System reell sein*** (vgl. Anm.[20])).

Beweis. Für einen solchen Operator gibt es eine abg. lin. M. $\mathfrak{M}$, die ihn reduziert, so daß sein in $\mathfrak{M}$ gelegener Teil $\bar{R}$ oder $-\bar{R}$ ist (Satz 38), hätte er nun eine der vier zuerst genannten Eigenschaften, so müßte dieselbe auch $\bar{R}$ oder $-\bar{R}$ zukommen. Aber $\bar{R}$ (und darum auch $-\bar{R}$) besitzt keine einzige derselben, wie sich bei seiner genauen Diskussion in Anhang IV zeigen wird (im § 4). Mit der letzten steht es so: Wenn ein Operator in einem vollst. norm. orth. System im Sinne von Anm.[20]) reell ist, so besteht für ihn kein Unterschied zwischen i und $-i$ — also sind seine Defektindizes einander gleich. Bei einem max. und nicht hypermax. H. O. sind sie aber verschieden (vgl. Satz 33). —

---

[57]) Sei $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System, wir zerlegen es in $m + n$ Teilfolgen $\psi_1^1, \psi_2^1, \ldots, \ldots, \psi_1^m, \psi_2^m, \ldots, \chi_1^1, \chi_2^1, \ldots, \ldots, \chi_1^n, \chi_2^n, \ldots$ (es kann auch m oder $n = \infty$ sein), und setzen $\mathfrak{M}_1, \ldots, \mathfrak{M}_m, \mathfrak{N}_1, \ldots, \mathfrak{N}_n$ den von einer jeden dieser Teilfolgen aufgespannten abg. lin. M. gleich.

Wir gehen nun zur systematischen Untersuchung der halbbeschränkten Operatoren über.

Satz 41. *Ein H.O. R, der alle Werte annimmt, ist hypermax. Er nimmt jeden Wert genau einmal an.*

Beweis. Sei f ein Erweiterungselement von R, ihm zugeordnet sei f^*. Es gibt nach Annahme ein g mit sinnvollem $Rg = f^*$. Für alle h mit sinnvollem Rh gilt dann: $(f, Rh) = (f^*, h) = (Rg, h) = (g, Rh)$, und da Rh alle Werte annimmt, muß $f = g$ sein; also ist Rf sinnvoll. Daher ist R hypermax., wir haben aber mitbewiesen: aus $Rf = Rg$ folgt, indem wir $f^* = Rf$ setzen, $f = g$.

Satz 42. *Ein lin. H.O. R, für den stets $(f, Rf) \geqq |f|^2$ gilt, kann zu einem definiten hypermax. H.O. fortgesetzt werden.*

Beweis. Zunächst bilden wir den abg. lin. H.O. $\tilde{R}$, auch für diesen gilt offenbar $(f, \tilde{R}f) \geqq |f|^2$. Der Wertevorrat von $\tilde{R}$ sei $\mathfrak{M}$, $\mathfrak{M}$ ist offenbar eine lin. M. Es ist

$$|f|^2 \leqq (f, \tilde{R}f) \leqq |f| \cdot |\tilde{R}f|, \quad |\tilde{R}f| \geqq |f|, \quad |\tilde{R}f - \tilde{R}g| = |\tilde{R}(f-g)| \geqq |f-g|,$$

also folgt aus $\tilde{R}f = \tilde{R}g$ $f = g$, d. h. $\tilde{R}$ hat eine Inverse $\tilde{R}^{-1}$. $\tilde{R}^{-1}$ ist offenbar auch abg. lin., in $\mathfrak{M}$ definiert, und wegen $|f-g| \geqq |\tilde{R}^{-1}f - \tilde{R}^{-1}g|$ stetig — also ist $\mathfrak{M}$ abg. Also können wir die abg. lin. M. $\mathfrak{N} = \mathfrak{H} - \mathfrak{M}$ bilden.

f gehöre zu $\mathfrak{N}$. Wenn $\tilde{R}g$ Sinn hat, so gehört es zu $\mathfrak{M}$, also ist $(f, \tilde{R}g) = 0$ — d. h. f ist Erweiterungselement, zu ihm gehört 0. Wenn also $\tilde{R}f$ Sinn hat, so ist $\tilde{R}f = 0$, $f = 0$. Der Definitionsbereich $\mathfrak{A}$ von $\tilde{R}$ (eine lin. M.!) und $\mathfrak{N}$ sind also lin. unabh. Wir können daher einen neuen Operator S so definieren: Sf hat Sinn, wenn $f = g + h$ (g von $\mathfrak{A}$, h von $\mathfrak{N}$, die Zerlegung ist ja eindeutig!) ist, und zwar $= \tilde{R}g$. S ist Forts. von $\tilde{R}$, und ein H.O., denn wenn Sf, Sf' Sinn haben ($f = g + h$, $f' = g' + h'$, g, g' von $\mathfrak{A}$, h, h' von $\mathfrak{N}$), so ist:

$$(f, Sg) = (g + h, \tilde{R}g') = (g, \tilde{R}g') = (\tilde{R}g, g') = (\tilde{R}g, g' + h') = (Sf, g).$$

$\mathfrak{N}$ reduziert S. In $\mathfrak{N}$ ist ja nach Definition Sf immer sinnvoll, und zwar $= 0$; da Sf stets einem $\tilde{R}g$ gleich, also in $\mathfrak{M}$ gelegen ist, haben wir auch $P_{\mathfrak{N}}Sf = 0 = SP_{\mathfrak{N}}f$. Also reduziert auch $\mathfrak{M} = \mathfrak{H} - \mathfrak{N}$ S, die in $\mathfrak{M}$, $\mathfrak{N}$ liegenden Teile von S seien S' bzw. S''. Der Wertevorrat von S ist $\mathfrak{M}$, wegen $Sf = P_{\mathfrak{M}}Sf = SP_{\mathfrak{M}}f$ nimmt S jeden seiner Werte schon in $\mathfrak{M}$ an — also hat S' auch den Wertevorrat $\mathfrak{M}$. Als ein-eindeutig lin. Bild des unendlichvieldimensionalen (weil überall dichten) $\mathfrak{A}$ ist es $\mathfrak{M}$ ebenfalls, also ist $\mathfrak{M}$ ein Hilbertscher Raum; darum ist Satz 41 anwendbar: S' ist hypermax., S'' ist sogar $= 0$.

Hieraus folgert man mühelos, daß S selbst hypermax. sein muß. Obendrein ist es definit, denn sei $f = g + h$, g von $\mathfrak{A}$, h von $\mathfrak{N}$:

$$(f, Sf) = (g + h, \tilde{R} g) = (g, \tilde{R} g) \geqq |g|^2 \geqq 0.$$

Damit ist alles bewiesen.

Satz 43. *Ein lin. H.O. R, für den stets $(f, Rf) \geqq C \cdot |f|^2$ oder stets $\leqq C \cdot |f|^2$ gilt (die Bedingungen der Halbbeschränktheit!) kann zu einem hypermax. H.O. S fortgesetzt werden, für den stets $(f, Sf) \geqq C' \cdot |f|^2$ bzw. stets $\leqq C' \cdot |f|^2$ gilt — dabei kann für C' jede Zahl $< C$ bzw. $> C$ gewählt werden*[58]).

Beweis. Für $C' = 0$, $C = 1$ (und $>$-Zeichen) ist dies Satz 42, und alle anderen Fälle können hierauf zurückgeführt werden, indem man $\frac{1}{C - C'} \cdot R - \frac{C'}{C - C'} \cdot 1$, $\frac{1}{C - C'} \cdot S - \frac{C'}{C - C'} \cdot 1$ statt R, S betrachtet.

XII. Pathologie der unbeschränkten Operatoren.

Zunächst einige Hilfssätze.

Satz 44. *Wenn R ein nach oben nicht halbbeschränkter lin. H.O. ist, so gibt es (für jedes $n = 1, 2, \ldots$ und jedes C) eine n-dimensionale lin. M., in der Rf stets Sinn hat, und durchweg $(f, Rf) \geqq C \cdot |f|^2$ ist.*

Beweis. Seien $f_1, \ldots, f_k$ irgendwelche Elemente, für die $Rf_1, \ldots, Rf_k$ sinnvoll ist. Wir zeigen zuerst: es muß ein φ geben, so daß φ zu allen $f_1, \ldots, f_k$ orth. ist, $|\varphi| = 1$, $R\varphi$ sinnvoll, und $(\varphi, R\varphi) \geqq D$. Nehmen wir nämlich das Gegenteil an. Da es nur auf die von den $f_1, \ldots, f_k$ aufgespannte lin. M. ankommt, können wir sie „orthogonalisieren" (vgl. Satz 8), d. h. durch die norm. orth. $\varphi_1, \ldots, \varphi_l$ $(l \leqq k)$ ersetzen. Sei f beliebig, aber mit sinnvollem Rf. Dann ist $f = \sum_{\mu=1}^{l} a_\mu \varphi_\mu + b\varphi$, wo $|\varphi| = 1$, φ zu allen $\varphi_1, \ldots, \varphi_l$ orth. ist (wir „orthogonalisieren" f hinzu), da $Rf, R\varphi_1, \ldots, R\varphi_l$ Sinn haben, hat auch $R\varphi$ Sinn — nach Annahme muß $(\varphi, R\varphi) \leqq D$ sein. Es ist

$$|f|^2 = \sum_{\mu=1}^{l} |a_\mu|^2 + |b|^2,$$

$$(f, Rf) = \sum_{\mu,\nu=1}^{l} a_\mu \bar{a}_\nu (\varphi_\mu, R\varphi_\nu) + 2\mathfrak{R} \sum_{\mu=1}^{l} b \bar{a}_\mu (\varphi, R\varphi_\mu) + |b|^2 (\varphi, R\varphi)$$

$$\leqq \sum_{\mu,\nu=1}^{l} |a_\mu| |a_\nu| \cdot |\varphi_\mu| |R\varphi_\nu| + 2 \sum_{\mu=1}^{l} |b| |a_\mu| \cdot |\varphi| |R\varphi_\mu| + |b|^2 \cdot D,$$

[58]) Der Satz gilt wohl auch noch für $C' = C$, es liegt aber kein Beweis dafür vor.

also wenn wir $D' = \operatorname*{Max}_{\mu=1,\ldots,l} (|R\varphi_\mu|, D)$ setzen,

$$\leqq D' \cdot \left(\sum_{\mu=1}^{l} |a_\mu| + |b|\right)^2 \leqq (l+1) D' \cdot \left(\sum_{\mu=1}^{l} |a_\mu|^2 + |b|^2\right) = (l+1) D' \cdot |f|^2;$$

also wäre R nach oben halbbeschränkt, entgegen der Annahme.

Wir gehen nun einen Schritt weiter. $g_1, \ldots, g_k$ seien nunmehr ganz willkürlich, es gibt dann ein φ mit sinnvollem $R\varphi$, $|\varphi| = 1$, $(\varphi, R\varphi) \geqq D$, so daß alle $|\varphi, g_\mu)| \leqq \varepsilon$ sind ($\mu = 1, \ldots, k$, $\varepsilon > 0$). (Also nicht ganz genau orth.!) Die f mit sinnvollem Rf liegen ja überall dicht, wir wählen k solche $f_1, \ldots, f_k$ mit $|g_\mu - f_\mu| \leqq \varepsilon$ ($\mu = 1, \ldots, k$), und φ nach dem vorher Bewiesenen. Dann ist $|\varphi| = 1$, $(\varphi, R\varphi) \geqq D$, und

$$|(\varphi, g_\mu)| = |(\varphi, g_\mu - f_\mu)| \leqq |\varphi| \cdot |g_\mu - f_\mu| \leqq \varepsilon.$$

Jetzt können wir die eigentliche Behauptung beweisen. Wir wählen nun sukzessiv: φ_1 mit sinnvollem $R\varphi_1$, $|\varphi_1| = 1$, $(\varphi_1, R\varphi_1) \geqq D$; φ_2 mit sinnvollem $R\varphi_2$, $|\varphi_2| = 1$, $(\varphi_2, R\varphi_2) \geqq D$ und $|(R\varphi_1, \varphi_2)| \leqq 1; \ldots;$ φ_n mit sinnvollem $R\varphi_n$, $|\varphi_n| = 1$, $(\varphi_n, R\varphi_n) \geqq D$ und $|(R\varphi_1, \varphi_n)| \leqq 1, \ldots,$ $|(R\varphi_{n-1}, \varphi_n)| \leqq 1$. Also: alle $R\varphi_\mu$ haben Sinn, es ist $|\varphi_\mu| = 1$, $(\varphi_\mu, R\varphi_\mu) \geqq D$, und $|(R\varphi_\mu, \varphi_\nu)| \leqq 1$ für $\mu < \nu$, also immer für $\mu \neq \nu$. Daraus folgt:

$$\left|\sum_{\mu=1}^{n} a_\mu \varphi_\mu\right|^2 = \sum_{\mu,\nu=1}^{n} a_\mu \bar{a}_\nu (\varphi_\mu, \varphi_\nu) \leqq \sum_{\mu,\nu=1}^{n} |a_\mu| |a_\nu| \cdot |\varphi_\mu| |\varphi_\nu|$$

$$\leqq \left(\sum_{\mu=1}^{n} |a_\mu|\right)^2 \leqq n \cdot \sum_{\mu=1}^{n} |a_\mu|^2,$$

$$\left(\sum_{\mu=1}^{n} a_\mu \varphi_\mu, R\left\{\sum_{\mu=1}^{n} a_\mu \varphi_\mu\right\}\right) = \sum_{\mu,\nu=1}^{n} a_\mu \bar{a}_\nu (\varphi_\mu, R\varphi_\nu)$$

$$= \sum_{\mu=1}^{n} |a_\mu|^2 (\varphi_\mu, R\varphi_\mu) + \sum_{\substack{\mu,\nu=1 \\ (\mu \neq \nu)}}^{n} a_\mu \bar{a}_\nu (\varphi_\mu, R\varphi_\nu)$$

$$\geqq D \sum_{\mu=1}^{n} |a_\mu|^2 - \sum_{\substack{\mu,\nu=1 \\ (\mu \neq \nu)}}^{n} |a_\mu| |a_\nu| = (D+1) \sum_{\mu=1}^{n} |a_\mu|^2 - \left(\sum_{\mu=1}^{n} |a_\mu|\right)^2$$

$$\geqq (D + 1 - n) \sum_{\mu=1}^{n} |a_\mu|^2.$$

Für $D > n - 1$ hat also die zweite Ungleichung die Konsequenz, daß $\sum_{\mu=1}^{n} a_\mu \varphi_\mu = 0$ auch $\sum_{\mu=1}^{n} |a_\mu|^2 = 0$, $a_1 = \ldots = a_n = 0$ nach sich zieht — d. h. die $\varphi_1, \ldots, \varphi_n$ sind lin. unabh., die lin. M. der $\sum_{\mu=1}^{n} a_\mu \varphi_\mu$ ist n-dimensional. Weiter ist (nach beiden Ungleichungen)

$$\left(\sum_{\mu=1}^{n} a_\mu \varphi_\mu, R\left\{\sum_{\mu=1}^{n} a_\mu \varphi_\mu\right\}\right) \geqq \frac{D+1-n}{n} \left|\sum_{\mu=1}^{n} a_\mu \varphi_\mu\right|^2,$$

also gilt, wenn wir noch $D \geqq nC + n - 1$ nehmen, in dieser lin. M. durchweg $(f, Rf) \geqq C \cdot |f|^2$. Damit sind wir am Ziele.

Satz 45. *Wenn R ein nach oben nicht halbbeschränkter lin. H. O. ist, und $f_1, \ldots, f_k$ irgendwelche Elemente, C eine beliebige Zahl, so gibt es ein φ mit sinnvollem $R\varphi$, das zu allen $f_1, \ldots, f_k$ orth. ist, und für welches $|\varphi| = 1$, $(\varphi, R\varphi) \geqq C$ gilt.*

Beweis[59]). Wir wählen die lin. M. des Satzes 44 $k+1$-dimensional, in ihr sind dann die (höchstens k unabhängige lin. Bedingungen darstellenden) Gleichungen $(f_1, \varphi) = 0, \ldots, (f_k, \varphi) = 0$ durch ein $f \neq 0$ lösbar — durch Multiplizieren mit einer Zahl können wir $|f| = 1$ erreichen. Dieses f leistet alles Gewünschte.

Satz 46. *Ein nach oben nicht halbbeschränkter lin. H. O. R ist immer Forts. von einem definiten lin. H. O. S.*

Beweis. Es genügt eine überall dichte Folge $f_1, f_2, \ldots$ anzugeben, so daß alle Rf_μ sinnvoll sind und stets

$$\left(\sum_{\mu=1}^{n} x_\mu f_\mu, R\left\{\sum_{\mu=1}^{n} x_\mu f_\mu\right\}\right) \geqq 0 \qquad (n = 1, 2, \ldots)$$

gilt. Dann ist nämlich der in der von $f_1, f_2, \ldots$ aufgespannten lin. M. (überall dicht, also nicht abg.!) definierte und dort mit R übereinstimmende Operator S offenbar ein lin. H. O. und definit — und R ist Forts. von ihm.

Sei $g_1, g_2, \ldots$ eine überall dichte Folge, so daß alle Rg_μ sinnvoll sind (d. h. $g_1, g_2, \ldots$ sei eine im Definitionsbereiche von R, $\mathfrak{A}$, überall dichte Teilfolge von $\mathfrak{A}$, $\mathfrak{A}$ ist ja seinerseits überall dicht; so etwas gibt es, weil $\mathfrak{A}$ als Teil von $\mathfrak{H}$ separabel ist, vgl. Anm. [33])), und $C_1, C_2, \ldots$ eine Folge von Zahlen, über die wir noch näher verfügen werden. Wir wählen nun eine Folge $\varphi_1, \varphi_2, \ldots$ auf Grund von Satz 45 mit den folgenden Eigenschaften aus: $R\varphi_1$ sinnvoll, $|\varphi| = 1$, $(\varphi_1, R\varphi_1) \geqq C_1$; $R\varphi_2$ sinnvoll, $|\varphi^2| = 1$, $(\varphi_2, R\varphi_2) \geqq C_2$ und $(R\varphi_1, \varphi_2) = 0$: $R\varphi_3$ sinnvoll, $|\varphi_3| = 1$, $(\varphi_3, R\varphi_3) \geqq C_3$ und $(R\varphi_1, \varphi_3) = (R\varphi_2, \varphi_3) = 0$; .... Somit sind alle $R\varphi_\mu$ sinnvoll, $|\varphi_\mu| = 1$, $(\varphi_\mu, R\varphi_\mu) \geqq C_\mu$ und $(R\varphi_\mu, \varphi_\nu) = 0$ für $\mu < \nu$, also immer für $\mu \neq \nu$. Wir setzen $f_1 = g_1 + \varphi_1$, $f_2 = g_2 + \frac{1}{2}\varphi_2$, $f_3 = g_3 + \frac{1}{3}\varphi_3, \ldots$, und behaupten, daß die Folge $f_1, f_2, \ldots$ (bei geeigneter Wahl der $C_1, C_2, \ldots$) alles Gewünschte leistet.

Zunächst ist sie überall dicht: $|f_n - g_n| = \left|\frac{1}{n}\varphi_n\right| = \frac{1}{n} \to 0$, zu jedem f gibt es eine Teilfolge $g_{N_1}, g_{N_2}, \ldots$ von $g_1, g_2, \ldots$ mit $g_{N_n} \to f$, dann ist auch $f_{N_n} \to f$. Nun berechnen wir (f, Rf) für die Linearaggregate der $f_1, f_2, \ldots$:

---

[59]) Bei sinnvollen $Rf_1, \ldots, Rf_k$ hatten wir dies schon am Anfange des Beweises von Satz 44, aber wir müssen uns von dieser Annahme befreien.

$$\Big(\sum_{\mu=1}^{n} a_\mu f_\mu,\ R\Big\{\sum_{\mu=1}^{n} a_\mu f_\mu\Big\}\Big)=\Big(\sum_{\mu=1}^{n} a_\mu g_\mu+\sum_{\mu=1}^{n}\frac{a_\mu}{\mu}\varphi_\mu,\ \sum_{\mu=1}^{n} a_\mu R g_\mu+\sum_{\mu=1}^{n}\frac{a_\mu}{\mu}R\varphi_\mu\Big)$$

$$=\sum_{\mu,\nu=1}^{n} a_\mu \bar a_\nu (g_\mu, R g_\nu)+2\Re\sum_{\mu,\nu=1}^{n}\frac{a_\mu \bar a_\nu}{\mu}(\varphi_\mu, R g_\nu)+\sum_{\mu,\nu=1}^{n}\frac{|a_\mu|^2}{\mu^2}(\varphi_\mu, R\varphi_\mu)$$

$$+\sum_{\substack{\mu,\nu=1\\(\mu\neq\nu)}}^{n}\frac{a_\mu \bar a_\nu}{\mu\nu}(\varphi_\mu, R\varphi_\nu)$$

$$\geqq-\sum_{\mu,\nu=1}^{n}|a_\mu||a_\nu|\cdot|g_\mu||R g_\nu|-2\sum_{\mu,\nu=1}^{n}|a_\mu||a_\nu|\cdot\frac{|\varphi_\mu||R g_\nu|}{\nu}+\sum_{\mu=1}^{n}|a_\mu|^2\cdot\frac{C_\mu}{\mu^2}$$

$$=\sum_{\mu=1}^{n}\frac{C_\mu}{\mu^2}\cdot|a_\mu|^2-\sum_{\mu,\nu=1}^{n}\Big(|g_\mu||R g_\nu|+\frac{|R g_\nu|}{\nu}\Big)|a_\mu||a_\nu|.$$

Wir setzen $|g_\mu||R g_\nu|+\frac{|R g_\nu|}{\nu}=\bar A_{\mu\nu}$ (dies hängt von $g_1, g_2, \dots$ ab, aber nicht von den $f_1, f_2, \dots$ cder $\varphi_1, \varphi_2, \dots$, d. h. den $C_1, C_2, \dots$!) und schätzen den Subtrahenden ab:

$$\sum_{\mu,\nu=1}^{n}\bar A_{\mu\nu}|a_\mu||a_\nu|=\sum_{\mu=1}^{n}\bar A_{\mu\mu}|a_\mu|^2+\sum_{\substack{\mu,\nu=1\\(\mu<\nu)}}^{n}(\bar A_{\mu\nu}+\bar A_{\nu\mu})|a_\mu||a_\nu|$$

$$\leqq\sum_{\mu=1}^{n}\bar A_{\mu\mu}|a_\mu|^2+\sum_{\substack{\mu,\nu=1\\(\mu<\nu)}}^{n}\Big(2^{\nu-\mu-1}(\bar A_{\mu\nu}+\bar A_{\nu\mu})|a_\nu|^2+\frac{1}{2^{\nu-\mu}}|a_\mu|^2\Big)$$

$$=\sum_{\mu=1}^{n}\Big\{\bar A_{\mu\mu}+\sum_{\varrho=1}^{\mu-1}2^{\mu-\varrho-1}(\bar A_{\mu\varrho}+\bar A_{\varrho\mu})+\sum_{\varrho=\mu+1}^{n}\frac{1}{2^{\varrho-\mu}}\Big\}|a_\mu|^2$$

$$\leqq\sum_{\mu=1}^{n}\Big\{\bar A_{\mu\mu}+1+\sum_{\sigma=0}^{\mu-2}2^{\sigma}(\bar A_{\mu,\mu-\sigma-1}+\bar A_{\mu-\sigma-1,\mu})\Big\}|a_\mu|^2.$$

Indem wir den $\{\ \}$-Ausdruck kurz mit $\bar B_\mu$ bezeichnen, wird:

$$\Big(\sum_{\mu=1}^{n} a_\mu f_\mu,\ R\Big\{\sum_{\mu=1}^{n} a_\mu f_\mu\Big\}\Big)\geqq\sum_{\mu=1}^{n}\Big(\frac{1}{\mu^2}C_\mu-\bar B_\mu\Big)\cdot|a_\mu|^2.$$

Es genügt also $C_\mu=\mu^2\cdot\bar B_\mu$ zu setzen und wir sind am Ziele [60]. —

---

[60]) Für die Zwecke der in Anm. [22]) angekündigten Arbeit soll hier noch eine gewisse Verschärfung dieser Abschätzung ausgeführt werden (hier brauchen wir sie nicht). Es ist

$$\Big|\sum_{\mu=1}^{n} a_\mu f_\mu\Big|^2=\Big|\sum_{\mu=1}^{n} a_\mu g_\mu+\sum_{\mu=1}^{n}\frac{a_\mu}{\mu}\varphi_\mu\Big|^2$$

$$=\sum_{\mu,\nu=1}^{n} a_\mu \bar a_\nu (g_\mu, g_\mu)+2\Re\sum_{\mu,\nu=1}^{n}\frac{a_\mu \bar a_\nu}{\mu}(\varphi_\mu, g_\nu)+\sum_{\mu,\nu=1}^{n}\frac{a_\mu \bar a_\nu}{\mu\nu}(\varphi_\mu, \varphi_\nu)$$

$$\leqq\sum_{\mu,\nu=1}^{n}\Big\{|g_\mu||g_\nu|+\frac{|\varphi_\mu||g_\nu|}{\mu}+\frac{|\varphi_\mu||\varphi_\nu|}{\mu\nu}\Big\}|a_\mu||a_\nu|$$

$$=\sum_{\mu,\nu=1}^{n}\Big\{|g_\mu||g_\nu|+\frac{|g_\nu|}{\mu}+\frac{1}{\mu\nu}\Big\}|a_\mu||a_\nu|.$$

(Fortsetzung der Fußnote [60]) auf nächster Seite.)

Damit haben wir bereits ein recht „pathologisches“ Resultat gewonnen, das sich leicht weiter verschärfen läßt.

Satz 47. *Sei R ein nach oben oder nach unten nicht halbbeschränkter lin. H.O., dann gibt es einen hypermax. H.O. S, für den stets $(f, Sf) \geqq C \cdot |f|^2$ bzw. stets $\leqq C \cdot |f|^2$ gilt, so daß R, S beide Forts. desselben lin. H. O. sind — dabei kann für C jede Zahl gewählt werden.*

Beweis. Für $C = -1$ und nicht-Halbbeschränktheit nach oben folgt dies aus Satz 46 und Satz 43 (wo $C = 0$, $C' = -1$ zu setzen ist), für alle anderen C folgt es hieraus, wenn man $R - (C+1) \cdot 1$, $S - (C+1) \cdot 1$ bzw. $-R + (C-1) \cdot 1$, $-S + (C-1) \cdot 1$ statt R, S betrachtet. —

Hieraus folgen die sonderbarsten Dinge für unbeschränkte H. O. Das Verhältnis eines H. O. zu seiner Matrix und dieser zu ihren Abschnitten wird dadurch qualitativ anders wie bei beschränkten; insbesondere zeigt sich, daß die „Abschnittsspektra“ bei unbeschränkten Matrizen sehr wenig mit dem Spektrum der Matrix selbst zu tun haben und diese (nicht wie in der Hilbertschen Theorie) in keinem Sinne „approximieren“. Über das Verhältnis Operator-Matrix werden wir einiges im Anhang III sagen, die genaue Erörterung der Eigenschaften unbeschränkter Matrizen soll in der in Anm. [22]) angekündigten Arbeit erfolgen.

Wir geben noch einige charakteristische Eigenschaften der unbeschränkten H. O. an. Im Gegensatze zu Satz 41 bis 47 beschränken wir uns jetzt auf abg. lin. Operatoren.

Satz 48. *Sei R ein abg. lin. H.O. Dann sind die vier Aussagen α) R ist beschränkt, β) R ist beschränkt hypermax., γ) R ist überall*

Wir nennen den $\{\ \}$-Ausdruck kurz $\overline{C}_{\mu\nu}$, dann wird wie im Text:

$$\sum_{\mu,\nu=1}^{n} \overline{C}_{\mu\nu} |a_\mu| |a_\nu| \leqq \sum_{\mu=1}^{n} \Big\{ \overline{C}_{\mu\mu} + 1 + \sum_{\sigma=0}^{\mu-2} 2^{\sigma} (\overline{C}_{\mu,\mu-\sigma-1} + \overline{C}_{\mu-\sigma-1,\mu}) \Big\} |a_\mu|^2.$$

Diesen $\{\ \}$-Ausdruck nennen wir $\overline{D}_\mu$, dann wird

$$\Big| \sum_{\mu=1}^{n} a_\mu f_\mu \Big|^2 \leqq \sum_{\mu=1}^{n} \overline{D}_\mu |a_\mu|^2.$$

Wenn wir nun $C_\mu = \mu^2 (\overline{B}_\mu + \mu \overline{D}_\mu)$ setzen, so ist für $a_1 = \ldots = a_{p-1} = 0$ ($n \geqq p$!)

$$\Big(\sum_{\mu=p}^{n} a_\mu f_\mu, R\Big\{ \sum_{\mu=p}^{n} a_\mu f_\mu \Big\}\Big) - p \Big| \sum_{\mu=p}^{n} a_\mu f_\mu \Big|^2 \geqq \sum_{\mu=p}^{n} \Big(\frac{C_\mu}{\mu^2} - \overline{B}_\mu - p \overline{D}_\mu \Big) |a_\mu|^2 \geqq 0.$$

D. h.: in der von den $f_p, f_{p+1}, f_{p+2}, \ldots$ aufgespannten lin. M. (auch diese ist natürlich überall dicht, also nicht abg.!) gilt sogar

$$(f, Rf) \geqq p \cdot |f|^2,$$

für jedes $p = 1, 2, \ldots$.

sinnvoll, δ) es gibt keinen abg. lin. H. O. von dem R eig. Forts. ist, alle gleichwertig[61]).

Beweis. Wenn R nach oben nicht halbbeschränkt ist, so wählen wir C so, daß $(f, Rf) < C \cdot |f|^2$ vorkommt, und dann einen hypermax. H. O. S nach Satz 47, so daß stets $(f, Rf) \geqq C \cdot |f|^2$ gilt — dann kann S offenbar keine Forts. von R sein. Nun sind R, S Forts. eines lin. H. O. T, und weil sie abg. lin. sind, auch des abg. lin. H. O. $\tilde{T}$. Da S nicht Forts. von R ist, muß $\tilde{T} \neq R$ sein, d. h. δ) ist nicht erfüllt. Wenn R nach unten nicht halbbeschränkt ist, so gilt δ) nicht für $-R$, also auch nicht für R. Also folgt allenfalls aus der Unbeschränktheit das Gegenteil von δ), d. h. aus δ) folgt α).

Aus α) folgt γ): denn R ist abg. und stetig, hat also einen abg. Definitionsbereich; da aber dieser überall dicht ist, muß er $= \mathfrak{H}$ sein. Aus γ) folgt δ): denn wäre R eig. Forts. des abg. lin. H. O. S, so wäre dieser nicht max., also zu mehreren max. H. O. festsetzbar (Zusatz zu Satz 35), z. B. zu $R' \neq R$; dann ist für alle f mit sinnvollen $R'f$ (wenn Sg Sinn hat)

$$(R'f, g) = (f, R'g) = (f, Sg) = (f, Rg) = (Rf, g),$$

und weil die Menge dieser g überall dicht ist, $R'f = Rf$ — also wäre R Forts. von R', $R' \neq R$, trotzdem R' max. sein sollte.

Wir haben das logische Schema

$$\alpha) \rightarrow \gamma) \rightarrow \delta) \rightarrow \alpha),$$

d. h. α), γ), δ) sind logisch gleichwertig. Aus β) folgt α), und β) folgt aus α) und γ): daher ist es ihnen auch gleichwertig.

Satz 49. *Sei R ein unbeschränkter abg. lin. H. O., dann ist es Forts. eines abg. lin. H. O., der zu Kontinuum vielen verschiedenen max. H. O. fortsetzbar ist.*

Beweis. Nach Satz 48 δ) muß R eig. Forts. eines abg. lin. H. O. S sein, dieses S ist nicht max., hat also nach dem Zusatz zu Satz 35 die gewünschte Eigenschaft.

Anhang I. Funktionenräume.

Hier soll der im Kap. I angekündigte Nachweis erbracht werden, daß alle in der Einleitung I. und II. erwähnten Funktionenräume abstrakte Hilbertsche Räume sind, d. h. den Bedingungen A bis E aus Kap. I genügen. Für den „gewöhnlichen Hilbertschen Raum“ $\left(\text{aller Folgen } (x_1, x_2, \ldots) \text{ mit } \sum_{n=1}^{\infty} |x_n|^2\right)$ erübrigt sich dieser Beweis, sobald mindestens ein A bis E ge-

[61]) Die Gleichwertigkeit von α) und γ) ist der bekannte Satz von Toeplitz, Math. Annalen **69** (1911), S. 321.

nügendes System bekannt ist — denn jedes solche System ist ja (nach den Schlußüberlegungen von Kap. I) dem „gewöhnlichen Hilbertschen Raume“ isomorph, also muß dieser die Eigenschaften A bis E gleichfalls besitzen. Wir müssen daher nur noch die Funktionenräume von Einleitung I. betrachten.

Ω sei also einer der dort genannten metrischen Räume, oder auch irgendein allgemeineres Gebilde, von dem wir aber jedenfalls folgendes verlangen: Es soll in Ω einen Begriff der „Meßbarkeit“, ein „Maß“ und ein „Integral“ ($\int\limits_{\Omega} f(P)\,dv$ möge es heißen) geben, und zwar im Lebesgueschen Sinne. Wir wollen hier nicht näher ausführen, was dieser „Lebesguesche Sinn“ genau bedeuten soll, dies wäre leicht zu beschreiben, aber etwas weitläufig. Bei den nun auszuführenden Überlegungen wird man ohnehin erkennen, auf welche Eigenschaften dieser Begriffe es ankommt. Ferner setzen wir Ω auch als separabel voraus.

Der Funktionenraum $\mathfrak{H}'$ bestehe nunmehr aus allen in Ω definierten meßbaren (komplexwertigen) Funktionen $f(P)$, für die $\int\limits_{\Omega} |f(P)|^2\,dv$ endlich ist: dabei gelten zwei Funktionen, die nur auf einer Menge vom Maße 0 verschiedene Werte haben, als nicht verschieden. Was af und $f+g$ bedeuten, ist klar, (f, g) sei gleich $\int\limits_{\Omega} f(P)\overline{g(P)}\,dv$. Diese Definitionen sind zulässig: wenn f, g zu $\mathfrak{H}'$ gehören, also $\int\limits_{\Omega} |f(P)|^2\,dv$, $\int\limits_{\Omega} |g(P)|^2\,dv$ endlich sind, so gehören auch $af, f+g$ zu $\mathfrak{H}'$, denn $\int\limits_{\Omega} |af(P)|^2\,dv$ ist offenbar endlich, und $\int\limits_{\Omega} |f(P)+g(P)|^2\,dv$ ist es wegen $|f(P)+g(P)|^2 \leqq 2|f(P)|^2 + 2|g(P)|^2$; ferner ist $\int\limits_{\Omega} f(P)\overline{g(P)}\,dv$ sinnvoll, ja absolut konvergent, wegen $|f(P)\overline{g(P)}| \leqq \frac{1}{2}|f(P)|^2 + \frac{1}{2}|g(P)|^2$.

Wir gehen nun die Bedingungen A bis E der Reihe nach durch.

A, B sind offenbar erfüllt.

C. Es gilt eine Folge $f_1(P), f_2(P), \ldots$ zu finden, die im Sinne der Entfernung

$$|f-g| = \sqrt{\int\limits_{\Omega} |f(P)-g(P)|^2\,dv}$$

überall dicht ist. — Das Maß von Ω ist eventuell unendlich, jedenfalls können wir aber Ω durch eine aufsteigende Folge $\Pi_1, \Pi_2, \ldots$ von Teilmengen endlichen Maßes erschöpfen. Wenn f irgendeine Funktion aus $\mathfrak{H}'$ ist, so definieren wir f_N $(N = 1, 2, \ldots)$ so:

$$f_N(P) = \begin{cases} f(P), & \text{wenn } P \text{ in } \Pi_N \text{ liegt und } |f(P)| \leqq N \text{ ist,} \\ 0 & \text{sonst} \end{cases}$$

Da für jedes N $f_N(P) \to f(P)$ und alle $\int\limits_\Omega |f_N(P)|^2 dv$ beschränkt ($\leqq \int\limits_\Omega |f(P)|^2 dv$) sind, so muß $\int\limits_\Omega |f_N(P) - f(P)|^2 dv \to 0$, d. h. im Sinne unserer Entfernung $f_N \to f$. Also sind diejenigen f, die nur auf einer Menge endlichen Maßes Werte $\neq 0$ haben und die außerdem beschränkt sind, überall dicht — es genügt somit eine in ihnen überall dichte Folge zu finden.

Sei f eine Funktion dieser Klasse, das Maß der Menge wo sie $\neq 0$ ist, M, die obere Grenze ihres Absolutwertes A. Wir wählen im Intervalle $-A$, A eine Einteilung $\varrho_1, \ldots, \varrho_k$ rationaler Zahlen, die dieses mit einer Maschenweite $\leqq \varepsilon$ überdecken ($\varepsilon > 0$, sonst beliebig). Es sei

$$f_\varepsilon(P) = \begin{cases} \varrho_h, & \text{wenn } f(P) \neq 0 \text{ und } \varrho_h \text{ das } f(P) \text{ nächstgelegene der } \varrho_1, \ldots, \varrho_k \text{ ist,} \\ 0 & \text{sonst} \end{cases}$$

Hieraus folgt sofort $\int\limits_\Omega |f_\varepsilon(P) - f(P)|^2 dv \leqq M \cdot 2A \cdot \varepsilon$, für $\varepsilon \to 0$ gilt somit für diese f_ε $f_\varepsilon \to f$. D. h. diejenigen Funktionen, die stets rationale Werte annehmen, und nur endlich viele verschiedene, und jeden der $\neq 0$ ist nur auf einer Menge endlichen Maßes, sind auch überall dicht. Die gesuchte Folge braucht daher nur in diesen überall dicht zu sein.

Wenn Σ irgendeine Teilmenge von endlichem Maße von Ω ist, so sei f_Σ die Funktion, die in Σ $=1$ und sonst $=0$ ist; die Funktionenklasse von vorhin ist einfach die aller $\varrho_1 f_{\Sigma_1} + \ldots + \varrho_k f_{\Sigma_k}$ ($\varrho_1, \ldots, \varrho_k$ rational). Wenn wir eine Folge von Mengen $\Sigma^{(1)}, \Sigma^{(2)}, \ldots$ fänden, die in der Gesamtheit aller Mengen endlichen Maßes überall dicht ist (die Entfernung von zwei Mengen Σ', Σ'' ist die ihrer Funktionen $f_{\Sigma'}, f_{\Sigma''}$, d. h. das Maß ihrer „Unterschiedsmenge" $(\Sigma' + \Sigma'') - (\Sigma' \cdot \Sigma'')$ [62])), so lägen die Funktionen $\varrho_1 f_{\Sigma^{(n_1)}} + \ldots + \varrho_k f_{\Sigma^{(n_k)}}$ ($k = 1, 2, \ldots$, $\varrho_1, \ldots, \varrho_k$ rational, $(n_1, \ldots, n_k = 1, 2, \ldots)$ in $\mathfrak{H}'$ überall dicht: d. h. wir wären am Ziele. Eine solche Mengenfolge $\Sigma^{(1)}, \Sigma^{(2)}, \ldots$ gilt es also zu finden.

Sei Σ eine Menge vom endlichen Maß M, bekanntlich ist M die untere Grenze der Maße aller Σ enthaltenden offenen Punktmengen; sei Σ' eine solche Menge mit einem Maße $\leqq M + \varepsilon$, dann hat die „Unterschiedsmenge" $\Sigma' - \Sigma$ ein Maß $\leqq \varepsilon$. D. h. die offenen Mengen liegen überall dicht. Da Ω separabel ist, gibt es in Ω eine Folge offener Mengen $\Lambda_1, \Lambda_2, \ldots$, die ein „gleichwertiges Umgebungssystem" bilden, also so, daß jede offene Menge Σ' die Vereinigungsmenge aller in ihr enthaltenen Λ_n ist — etwa von $\Lambda_{n_1}, \Lambda_{n_2}, \ldots$. Das Maß von Σ' sei M', dann konvergieren die Maße der $\Lambda_{n_1}, \Lambda_{n_1} + \Lambda_{n_2}, \Lambda_{n_1} + \Lambda_{n_2} + \Lambda_{n_3}, \ldots$ gegen M'. Sei etwa das-

[62]) $+, -, \cdot$ sind hier die Zeichen für das Vereinigen, Abziehen und Durchschnitts-Bilden bei Mengen.

jenige von $A_{n_1} + \ldots + A_{n_k}$ bereits $\geq M' - \varepsilon$, dann hat die „Unterschiedsmenge“ $\Sigma' - (A_{n_1} + \ldots + A_{n_k})$ ein Maß $\leq \varepsilon$. D. h.: die $A_{n_1} + \ldots + A_{n_k}$ $(k = 1, 2, \ldots, n_1, \ldots, n_k = 1, 2, \ldots)$ liegen auch überall dicht[63]), und diese sind abzählbar viel, womit wir am Ziele sind.

D. Sei $k = 1, 2, \ldots$, wir wählen k paarweise punktfremde Mengen $K_1, \ldots, K_k$ in Ω aus, jede von endlichem Maße $\neq 0$. Die Funktion $f_h(P)$ sei $= 1$ in K_h und $= 0$ sonst $(h = 1, \ldots, k)$; dann sind die $f_1, \ldots, f_k$ offenbar. lin. unabh. (und sie gehören zu $\mathfrak{H}'$).

E. Sei $f_1, f_2, \ldots$ eine Funktionenfolge, die dem Cauchyschen Konvergenzkriterium genügt (zu jedem $\varepsilon > 0$ gibt es ein $N = N(\varepsilon)$, so daß für $m, n \geq N$ $\int\limits_{\Omega} |f_m(P) - f_n(P)|^2 dv \leq \varepsilon$ ist). Wir wählen (in Anlehnung an einen Beweis von Weyl) eine Folge $N_1 < N_2 < N_3 < \ldots$ mit $N_p \geq N\left(\frac{1}{8^p}\right)$ aus und schließen dann so. Es ist

$$\int\limits_{\Omega} |f_{N_{p+1}}(P) - f_{N_p}(P)|^2 dv \leq \frac{1}{8^p},$$

also hat diejenige Menge, auf der

$$|f_{N_{p+1}}(P) - f_{N_p}(P)| \geq \frac{1}{2^p}$$

gilt, ein Maß $\leq \frac{1}{2^p}$. Also gilt

$$|f_{N_{p+1}}(P) - f_{N_p}(P)| \leq \frac{1}{2^p}, \qquad |f_{N_{p+2}}(P) - f_{N_{p+1}}(P)| \leq \frac{1}{2^{p+1}}, \ldots$$

überall, mit Ausnahme einer Menge vom Maße $\leq \frac{1}{2^p} + \frac{1}{2^{p+1}} + \ldots = \frac{1}{2^{p-1}}$. Aber dort, wo die obigen Ungleichheiten gelten, hat die Folge $f_{N_1}(P)$, $f_{N_2}(P), \ldots$ einen Limes; also überall mit Ausnahme einer Menge vom Maße $\leq \frac{1}{2^{p-1}}$. Da dies für alle p gilt, hat die Ausnahmemenge das Maß 0.

Wir definieren nun: $f(P) = \lim f_{N_p}(P)$, soweit dieser Limes existiert, d. h. überall bis auf eine 0-Menge, und sonst $= 0$. Wenn nun $m \geq N(\varepsilon)$ ist, so gilt für fast alle p (nämlich sobald $N_p \geq N(\varepsilon)$ ist) $\int\limits_{\Omega} |f_m(P) - f_{N_p}(P)|^2 \cdot dv \leq \varepsilon$. Also dürfen wir $p \to \infty$ ausführen, und es wird: $\int\limits_{\Omega} |f_m(P) - f(P)|^2 dv \leq \varepsilon$. Somit gehört erstens $f_m - f$, also auch f, zu $\mathfrak{H}'$, und zweitens ist $f_m \to f$. D. h. alles ist bewiesen.

Anhang II. Das Eigenwertproblem unitärer Operatoren.

1. Wir betrachten im folgenden durchweg in ganz $\mathfrak{H}$ sinnvolle und beschränkte (stetige) lin. Operatoren R, auf die also Satz 12 α), β) an-

[63]) Wir nehmen natürlich nur die von endlichem Maße.

wendbar ist. Der Hermitesche Charakter hingegen wird nicht immer Voraussetzung sein. Bei festem f hat $L(g) = (f, Rg)$ die Eigenschaften

$$L(a_1 g_1 + \ldots + a_n g_n) = \bar{a}_1 L(g_n) + \ldots + \bar{a}_n L(g_n),$$
$$|L(g)| \leq C \cdot |f| |g| = D \cdot |g|,$$

so daß der Satz von F. Riesz (vgl. Anm. [52])) anwendbar ist: $L(g) = (f^*, g)$. Wir nennen das so bestimmte f^* $R^* f$, R^* ist ein überall sinnvoller und offenbar lin. Operator. Es hat die Eigenschaften

$$(f, Rg) = (R^* f, g), \qquad (Rf, g) = (f, R^* g)$$

(die erste war Definition, die zweite folgt durch Vertauschen von f, g und nehmen der komplex-Konjugierten). Insbesondere ist

$$|R^* f|^2 = (R^* f, R^* f) = (f, R R^* f) \leq C \cdot |f| |R^* f|, \qquad |R^* f| \leq C \cdot |f|,$$

d. h. auch R^* beschränkt[64]).

Man verifiziert sofort $R^{**} = R$, $(R \pm S)^* = R^* \pm S^*$, $(RS)^* = S^* R^*$,[65]) $(aR)^* = \bar{a} R^*$, $0^* = 0$, $1^* = 1$. Für H. O. ist $R = R^*$ charakteristisch (wenn sie beschränkt sind). Jeder unitäre Operator U gehört zu unserer Klasse (er ist wegen $|Uf| = |f|$ beschränkt), und ist in ihr durch $UU^* = U^* U = 1$, d. h. $U^{-1} = U^*$ gekennzeichnet[66]).

2. Jetzt sei R sogar H. O., wir nehmen zuerst sein Eigenwertproblem nach der Methode von F. Riesz in Angriff. Sei $p(x) = \sum_{\nu=1}^{n} a_\nu x^\nu$ ein Polynom (die a_ν reell), dann ist $p(R) = \sum_{\nu=0}^{n} a_\nu R^\nu$ $(R^0 = 1)$ ein H. O. unserer Klasse; das Addieren, Subtrahieren, Multiplizieren dieser $p(R)$ ist offenbar ebenso zu vollziehen, wie das der $p(x)$ (also sind alle vertauschbar). Wir wählen C mit $|Rf| \leq C \cdot |f|$ (für alle f).

Satz 1*. *Wenn für* $-C \leq x \leq C$ $p(x) \geq 0$ *ist, so ist* $p(R)$ *definit.*

Beweis. Sei $\mathfrak{M}$ irgendeine endlichvieldimensionale (also abg.) lin. M., etwa durch das norm. orth. System $\varphi_1, \ldots, \varphi_k$ aufgespannt. $R' = P_{\mathfrak{M}} R P_{\mathfrak{M}}$ ist ein H. O. mit $|R' f| = |P_{\mathfrak{M}} R P_{\mathfrak{M}} f| \leq |R P_{\mathfrak{M}} f| \leq C \cdot |P_{\mathfrak{M}} f| \leq C \cdot |f|$, d. h. von unserer Klasse, der durch $\mathfrak{M}$ offenbar reduziert wird (sein in

---

[64]) Bei Matrizen ist dieses * das Nehmen der konjugiert-transponierten, ebenso bei Integraloperator-Kernen, bei Differentialoperatoren die adjungierte.

[65]) Es handelt sich um überall sinnvolle Operatoren, daher sind die $+, -, \cdot$ ohne weiteres sinnvoll (vgl. Anm. [37])).

[66]) Wenn U unitär ist, so hat es eine Inverse U^{-1}, und es ist $(f, U^{-1} g) = (Uf, UU^{-1} g) = (Uf, g) = (f, U^* g)$ für alle f, g, also $U^{-1} = U^*$. Wenn $U^{-1} = U^*$ ist, so ist $(Uf, Ug) = (f, U^* Ug) = (f, g)$, also für $f = g$ $|Uf| = |f|$, so daß U längentreu ist; da es überall Sinn hat und alle Werte annimmt (U^{-1} hat ja überall Sinn!), ist es unitär.

$\mathfrak{H} - \mathfrak{M}$ liegender Teil ist 0!). Sein in $\mathfrak{M}$ liegender Teil ist also ein k-dimensionaler H. O., etwa mit den Eigenwerten $\lambda_1, \ldots, \lambda_k$; wegen $|R'f| \leqq C \cdot |f|$ sind alle $|\lambda_1|, \ldots, |\lambda_k| \leqq C$. Auch $p(R')$ wird durch $\mathfrak{M}$ reduziert, und sein in $\mathfrak{M}$ liegender Teil hat die Eigenwerte $p(\lambda_1), \ldots, p(\lambda_k)$ — also lauter Zahlen $\geqq 0$, d. h. er ist definit. Also: Wenn f in $\mathfrak{M}$ liegt so ist

$$(f, p(R')f) \geqq 0.$$

Sei nun f beliebig, wir wählen $\mathfrak{M}$ als die von $f, Rf, \ldots, R^n f$ aufgespannte (abg.) lin. M., dann ist f in $\mathfrak{M}$, $R'f = Rf, \ldots, R'^n f = R^n f$, also $p(R')f = p(R)f$. Also $(f, p(R)f) \geqq 0$; und da f ganz beliebig war, ist $p(R)$ definit.

Satz 2*. *Wenn für* $-C \leqq x \leqq C$ $|p(x)| \leqq \varepsilon$ *ist, so ist* $|p(R)f| \leqq \varepsilon \cdot |f|$.

Beweis. Satz 1* ist auf $\varepsilon \pm p(x)$ anwendbar, also $(f, \{\varepsilon \cdot 1 \pm p(R)\} f) \geqq 0$, $|(f, p(R)f)| \leqq \varepsilon \cdot |f|^2$. Nach Satz 12 (Gleichwertigkeit von γ) und α)!) bedeutet das $|p(R)f| \leqq \varepsilon \cdot |f|$.

Satz 3*. *Wenn* $P(x)$ *ein im Intervalle* $-C \leqq x \leqq C$ *stetige Funktion ist, und* $p_1(x), p_2(x), \ldots$ *eine dort gleichmäßig gegen* $P(x)$ *strebende Polynomfolge, so hat* $p_n(R)f$ *einen Limes* f^*, *der nur von* R, f *und* $P(x)$ *abhängt.*

Beweis. Klar nach Satz 2*. —

Durch $P(R)f = f^*$ definieren wir also einen überall sinnvollen Operator, der lin. H. O. ist, weil es die $p_n(R)$ sind, und beschränkt nach Satz 48 (α) und (γ), oder auch weil es die $p_n(x)$ gleichmäßig sind (also auch die $p_n(R)$, Satz 2*).

Satz 4*. *Die* $P(R)$ *sind ebenso zu addieren, subtrahieren, multiplizieren wie die* $P(x)$, *Satz* 1*, 2* *gilt für sie auch; wenn* $P(x)$ *ein Polynom ist, stimmt die alte und neue Definition überein.*

Beweis. Die letzte Behauptung folgt daraus, daß $p_1(x) = p_2(x) = \ldots = P(x)$ gesetzt werden kann; die übrigen gelten für die $p(x)$, also aus Stetigkeitsgründen auch für $P(x)$.

3. Nun biegen wir vom F. Rieszschen Wege ab.

Satz 5*. *Sei* $\mathfrak{M}_R$ *die Menge aller* f *mit* $Rf = 0$, $\mathfrak{N}_R$ *die Menge aller* Rf *und ihrer Häufungspunkte. Dann ist* $\mathfrak{M}_R = \mathfrak{H} - \mathfrak{N}_R$.

Beweis. Die Rf bilden eine lin. M., es genügt zu zeigen, daß $\mathfrak{M}_R$ aus allen zu dieser lin. M. orth. g besteht. Nun bedeutet das orth. stehen auf diese, $(g, Rf) = 0$ für alle f, d. h. $(Rg, f) = 0$ — also $Rg = 0$. Damit sind wir fertig.

Satz 6*. *Wenn* R, S *definit sind, so ist* $\mathfrak{M}_{R+S}$ *Teil von* $\mathfrak{M}_R$, *und* $\mathfrak{N}_R$ *Teil von* $\mathfrak{N}_{R+S}$.

Beweis. Aus $(R+S)g=0$ folgt $0 \leqq (g, Rg) \leqq (g, (R+S)g) = 0$, $(g, Rg) = 0$. Nun gilt für definite R

$$|(f, Rg)|^2 \leqq (f, Rf) \cdot (g, Rg)$$

(man beweist dies ebenso, wie die entsprechende Gleichung für $R = 1$ bei Satz 1), also hat $(g, Rg) = 0$ die Folge $(f, Rg) = 0$ (für alle f!), $Rg = 0$. Die obige Relation bedeutet daher: $\mathfrak{M}_{R+S}$ Teil von $\mathfrak{M}_R$. Nach Satz 5* ist also $\mathfrak{N}_R$ Teil von $\mathfrak{N}_{R+S}$.

Wir definieren:

$$f_a(x) = \operatorname{Max}(x-a, 0), \quad \mathfrak{M}(R, a) = \mathfrak{M}_{f_a(R)}, \quad \mathfrak{N}(R, a) = \mathfrak{N}_{f_a(R)}.$$

Satz 7*. *In* $\mathfrak{M}(R, a)$ *bzw.* $\mathfrak{N}(R, a)$ *gilt durchweg*

$$(f, Rf) \leqq a \cdot |f|^2 \quad \textit{bzw.} \quad \geqq a \cdot |f|^2.$$

Beweis. Sei $g_a(x) = \operatorname{Max}(-(x-a), 0)$, da $f_a(x) g_a(x) = 0$, $f_a(x) + g_a(x) = x - a$ ist, haben wir für $S' = f_a(R)$, $S'' = g_a(R)$ nach Satz 4*. $S'S'' = S''S' = 0$, $S' - S'' = R - a \cdot 1$. Da stets $f_a(x) \geqq 0$, $g_a(x) \geqq 0$ ist, sind S', S'' auch definit. In $\mathfrak{M}(R, a) = \mathfrak{M}_{S'}$ ist $S'f = 0$, also $Rf = af - S''f$, $(f, Rf) = a \cdot |f|^2 - (f, S''f) \leqq a \cdot |f|^2$. In $\mathfrak{N}(R, a) = \mathfrak{N}_{S'}$ befinden wir uns (nach Satz 5* und $S''S' = 0$) in $\mathfrak{M}_{S''}$, also ist $S''f = 0$, und $Rf = af + S'f$, $(f, Rf) = a \cdot |f|^2 + (f, S'f) \geqq a \cdot |f|^2$.

Satz 8*. *Wenn* $a \leqq b$ *ist, so ist* $\mathfrak{M}(R, a)$ *Teil von* $\mathfrak{M}(R, b)$ *und* $\mathfrak{N}(R, b)$ *Teil von* $\mathfrak{N}(R, a)$, *ferner ist für* $a < -C$ *bzw.* $> C$ $\mathfrak{M}(R, a) = (0)$ *bzw.* $= \mathfrak{H}$ *und* $\mathfrak{N}(R, a) = \mathfrak{H}$ *bzw.* $= (0)$.

Beweis. Die erste Behauptung folgt, da $f_a(x) \geqq 0$, $f_b(x) - f_a(x) \geqq 0$ ist, also $f_a(R)$, $f_b(R) - f_a(R)$ definit, aus Satz 6*. Die zweite beweisen wir so: Gehörte ein $f \neq 0$ für $a < -C$ bzw. $> C$ zu $\mathfrak{M}(R, a)$ bzw. $\mathfrak{N}(R, a)$, so wäre

$$(f, Rf) \leqq a \cdot |f|^2 < -C \cdot |f|^2 \quad \text{bzw.} \quad \geqq a \cdot |f|^2 > C \cdot |f|^2,$$

entgegen der Definition von C (Satz 12 (γ)), also ist $\mathfrak{M}(R, a)$ bzw. $\mathfrak{N}(R, a) = (0)$. Also $\mathfrak{N}(R, a)$ bzw. $\mathfrak{M}(R, a) = \mathfrak{H} - (0) = \mathfrak{H}$.

Satz 9*. *Es gilt ganz allgemein*

$$(Rf, g) = \int a \, d(P_{\mathfrak{M}(R, a)} f, g)$$

(*wobei das Integral über irgend zwei Grenzen* $< -C$, $> C$ *zu erstrecken ist*).

Beweis. Für $f = g$ folgt es aus Satz 7*, 8*[67]). Durch Einsetzen von

---

[67]) R ist mit $f_a(R)$ vertauschbar (Satz 4*), wird also durch $\mathfrak{M}(R, a)$ reduziert (aus $f_a(R) f = 0$ folgt $f_a(R) \cdot Rf = R \cdot f_a(R) f = 0$, und Satz 20), d. h. es ist mit $P_{\mathfrak{M}(R, a)}$ vertauschbar.

Nun sei $x_0 < x_1 < \ldots < x_p$, $x_0 < -C$, $x_p > C$, dann ist

(Fortsetzung der Fußnote [67]) auf nächster Seite.)

$\frac{f+g}{2}$, $\frac{f-g}{2}$ und Subtrahieren entsteht die Gleichung allgemein für die Realteile; durch Ersetzen von f, g durch $i\,f, g$ auch für die Imaginärteile. —

Satz 9* ist im wesentlichen das Hilbertsche Resultat für beschränkte H. O. Wir gehen weiter und betrachten zwei H. O. R, S unserer Klasse.

Satz 10*. *Wenn R, S vertauschbar sind, so sind es auch $P_{\mathfrak{M}(R,a)}$, $P_{\mathfrak{M}(S,b)}$.*

Beweis. R, S sind vertauschbar, also auch alle $P(R), Q(S)$ (Polynome offenbar, andere Funktionen aus Stetigkeitsgründen) — also insbesondere $f_a(R)$, $f_b(S)$. Daher reduziert $\mathfrak{M}(R,a)$ $f_b(S)$ (vgl. die Überlegung in Anm.[67])), und $\mathfrak{N}(R,a) = 1 - \mathfrak{M}(R,a)$ ebenfalls. Darum muß $\mathfrak{M}(S,b)$ so gebildet werden, daß man die f mit $f_b(S)f = 0$ zuerst in $\mathfrak{M}(R,a)$ und dann in $\mathfrak{N}(R,a)$ bildet, und dann beide abg. lin. M. addiert. Hieraus schließt man mühelos die Vertauschbarkeit von $P_{\mathfrak{M}(R,a)}$ und $P_{\mathfrak{M}(S,b)}$.

4. Jetzt wenden wir uns den sog. normalen Operatoren A unserer Klasse zu, die durch $AA^* = A^*A$ charakterisiert sind. Die H. O. $R = \frac{1}{2}(A + A^*)$, $S = \frac{1}{2i}(A - A^*)$ sind dann vertauschbar, und es ist $A = R + iS$, $A^* = R - iS$. Wir bilden nun die P. O.

$$E(A, a, b) = P_{\mathfrak{M}(R,a)} \cdot P_{\mathfrak{N}(S,b)}.$$

Es gilt der Satz:

Satz 11*.[68]) *A und die $E(A, a, b)$ erfüllen die folgenden Bedingungen:*

$$(Rf, f) = \sum_{\nu=1}^{p} (R(P_{\mathfrak{M}(R,x_\nu)} - P_{\mathfrak{M}(R,x_{\nu-1})}) f, f) = \sum_{\nu=1}^{p} (R\, P_{\mathfrak{M}(R,x_\nu) - \mathfrak{M}(R,x_{\nu-1})} f, f)$$

$$= \sum_{\nu=1}^{p} (R\, P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f, f) = \sum_{\nu=1}^{p} (R P_{\mathfrak{N}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f, P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f).$$

(Das letztere, weil R mit $P_{\mathfrak{M}(R,x_\nu)} - P_{\mathfrak{M}(R,x_{\nu-1})} = P_{\mathfrak{M}(R,x_a)\cdot\mathfrak{N}(R,x_{\nu-1})}$ vertauschbar ist, also $P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})}\, R\, P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} = R P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{M}(R,x_{\nu-1})}$ gilt.) Nach Satz 7* ist also der einzelne Addend in $\sum\limits_{\nu=1}^{p}$

$$\begin{cases} \leqq x_\nu & |P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f|^2, \\ \geqq x_{\nu-1} & |P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f|^2, \end{cases}$$

$$|P_{\mathfrak{M}(R,x_\nu)\cdot\mathfrak{N}(R,x_{\nu-1})} f|^2 = |P_{\mathfrak{M}(R,x_\nu) - \mathfrak{M}(R,x_{\nu-1})} f|^2 = |P_{\mathfrak{M}(R,x_\nu)} f|^2 - |P_{\mathfrak{M}(R,x_{\nu-1})} f|^2.$$

Nach Definition des Stieltjesschen Integrals bedeutet aber

$$(Rf, f) \begin{cases} \leqq \sum\limits_{\nu=1}^{p} x_\nu & (|P_{\mathfrak{M}(R,x_\nu)} f|^2 - |P_{\mathfrak{M}(R,x_{\nu-1})} f|^2), \\ \geqq \sum\limits_{\nu=1}^{p} x_{\nu-1} & (|P_{\mathfrak{M}(R,x_\nu)} f|^2 - |P_{\mathfrak{M}(R,x_{\nu-1})} f|^2), \end{cases}$$

das aus unseren Relationen folgt, die behauptete Gleichung

$$(Rf, f) = \int a\, d\,|P_{\mathfrak{M}(R,a)} f|^2 = \int a\, d\,(P_{\mathfrak{M}(R,a)} f, f).$$

[68]) Vgl. hierzu Einleitung IX und Anm. [26]).

α) $E(A, a, b)$, $E(A, a', b')$ sind vertauschbar und ihr Produkt ist $E(A, \mathrm{Min}(a, a'), \mathrm{Min}(b, b'))$.

β) Für genügend kleines a oder b ist $E(A, a, b) = 0$. Für genügend großes a oder b ist es von a bzw. b unabhängig. Für genügend großes a und b ist es $= 1$.

Und es gelten die Gleichungen:

$$(Af, g) = \iint (a + ib)\, d(E(A, a, b) f, g),$$
$$(A^* f, g) = \iint (a - ib)\, d(E(A, a, b) f, g).$$

(Die $\iint$ sind über ein hinreichend großes Rechteck zu erstrecken.)

Beweis. Vergegenwärtigt man sich die aus Satz $9^*, 10^*$ folgenden Gleichungen

$$(Rf, g) = \iint a\, d(E(A, a, b) f, g), \quad (Sf, g) = \iint b\, d(E(A, a, b) f, g),$$

so folgt alles unmittelbar aus Satz $8^*, 9^*, 10^*$ und der Definition der $E(A, a, b)$. —

Die unitären Operatoren U sind wegen $UU^* = U^* U = 1$ auch normal, Satz 11^* gilt also für sie auch. Ihre $E(U, a, b)$ haben aber noch weitergehende Eigenschaften und diese wollen wir jetzt herleiten.

Zunächst gilt allgemein:

$$\begin{aligned}(Af, E(A, a, b) g) &= \iint (a' + ib')\, d(E(A, a', b') f, E(A, a, b) g)\\ &= \iint (a' + ib')\, d(E(A, a, b)\, E(A, a', b') f, g)\\ &= \iint (a' + ib')\, d(E(A, \mathrm{Min}(a, a'), \mathrm{Min}(b, b')) f, g)\\ &= \iint^{a\,b} (a' + ib')\, d(E(A, a', b') f, g).\end{aligned}$$

$$\begin{aligned}(Af, Ag) &= \iint (a - ib)\, d(Af, E(A, a, b) g)\\ &= \iint (a - ib)\, d\left\{\iint^{a\,b} (a' + ib')\, d(E(A, a', b') f, g)\right\} = (\text{vgl. Anm.}^{55}))\\ &= \iint (a - ib)(a + ib)\, d(E(A, a, b) f, g)\\ &= \iint (a^2 + b^2)\, d(E(A, a, b) f, g).\end{aligned}$$

$$(f, g) = \iint d(E(A, a, b) f, g).$$

Für unitäre A ist also wegen $(f, g) = (Af, Ag)$

$$\iint (1 - a^2 - b^2)\, d(E(A, a, b) f, g) = 0.$$

Und wenn wir $f = g$ setzen: $\iint (1 - a^2 - b^2)\, d\,|E(A, a, b) f|^2 = 0$.

5. Wenn $\mathfrak{R}$ irgendein Rechteck $a' \leq x \leq a''$, $b' \leq y \leq b''$ ist (Ecken: $a', b'; a'', b'; a', b''; a'', b''$), so definieren wir einen P.O.

$$
\begin{aligned}
E(A, \mathfrak{R}) &= E(A, a'', b'') - E(A, a'', b') - E(A, a', b'') + E(A, a', b') \\
&= P_{\mathfrak{M}(R, a')} \cdot (1 - P_{\mathfrak{M}(R, a'')}) \cdot P_{\mathfrak{M}(S, b')} \cdot (1 - P_{\mathfrak{M}(S, b'')}) \\
&= P_{\mathfrak{M}(R, a') \mathfrak{N}(R, a'') \mathfrak{M}(S, b') \mathfrak{N}(S, b'')}.
\end{aligned}
$$

Möge nun $\mathfrak{R}$ keinen einzigen Punkt mit dem Einheitskreis gemeinsam haben, dann ist in $\mathfrak{R}$ stets $1 - a^2 - b^2 \geqq \varepsilon$ oder stets $\leqq -\varepsilon$ $(\varepsilon > 0)$, und wir haben für jedes f mit $E(A, \mathfrak{R}) f = f$ (d.h. aus $\mathfrak{M}(R, a') \mathfrak{N}(R, a'') \mathfrak{M}(S, b') \mathfrak{N}(S, b'')$)

$$
\begin{aligned}
&\left| \iint (1 - a^2 - b^2)\, d \,|E(A, a, b) f|^2 \right| \\
&= \left| \iint_{\mathfrak{R}} (1 - a^2 - b^2)\, d \,|E(A, a, b) f|^2 \right| \geqq \varepsilon \left| \iint_{\mathfrak{R}} d \,|E(A, a, b) f|^2 \right| \\
&= \varepsilon \left| \iint d \,|E(A, a, b) f|^2 \right| = \varepsilon \cdot |f|^2. \text{[69]}
\end{aligned}
$$

Also muß $f = 0$ sein, d.h. es ist $E(A, \mathfrak{R}) = 0$. — Wenn $\mathfrak{R}, \mathfrak{S}$ zwei punktfremde Rechtecke sind (die Ränder dürfen sich treffen), so ist $E(A, \mathfrak{R}) \cdot E(A, \mathfrak{S}) = 0$ [70]. Wenn die beiden Rechtecke $\mathfrak{R}, \mathfrak{S}$ zusammen genau ein Rechteck $\mathfrak{T}$ ausmachen, so ist $E(A, \mathfrak{R}) + E(A, \mathfrak{S}) = E(A, \mathfrak{T})$ [71].

Durch wiederholte Anwendung dieser Tatsachen folgert man nun: wenn $\mathfrak{U}$ eine Summe von endlich vielen Rechtecken ist, so ist erstens

$$E(A, \mathfrak{U}) = E(A, \mathfrak{R}_1) + \ldots + E(A, \mathfrak{R}_n)$$

($\mathfrak{R}_1, \ldots, \mathfrak{R}_n$ seien die genannten Rechtecke) ein P. O.; zweitens ist er für alle Zerlegungen $\mathfrak{R}_1, \ldots, \mathfrak{R}_n$ von $\mathfrak{U}$ derselbe; drittens ist, wenn $\mathfrak{U}, \mathfrak{V}$ zwei solche Mengen sind und höchstens Randpunkte gemein haben, $E(A, \mathfrak{U}) \cdot E(A, \mathfrak{V}) = 0$, $E(A, \mathfrak{U} + \mathfrak{V}) = E(A, \mathfrak{U}) + E(A, \mathfrak{V})$. Außerdem muß, wenn $\mathfrak{U}$ den Einheitskreis nicht trifft, $E(A, \mathfrak{U}) = 0$ sein (weil alle $E(A, \mathfrak{R}_1), \ldots, E(A, \mathfrak{R}_n) = 0$ sind); hieraus folgert man mühelos weiter: wenn $\mathfrak{U}, \mathfrak{V}$ dieselben Punkte mit dem Einheitskreis gemein haben (und keine Randecke auf ihm), so ist $E(A, \mathfrak{U}) = E(A, \mathfrak{V})$.

Jetzt sind wir in der Lage, das Stieltjessche Integral $\iint$ sozusagen auf Polarkoordinaten zu transformieren.

---

[69]) Da f in $\mathfrak{M}(R, a')$ und $\mathfrak{N}(R, a'')$ liegt, ist für $a \leqq a'$ oder $\geqq a''$ $P_{\mathfrak{M}(R, a)} f = 0$ bzw. $= f$, also $E(A, a, b) f$ von a unabhängig; ebenso ist es für $b \leqq b'$ oder $\geqq b''$ von b unabhängig. Daher dürfen wir $\iint$ und $\iint_{\mathfrak{R}}$ identifizieren.

[70]) Seien $\mathfrak{R}, \mathfrak{S}$ etwa $a' \leqq x \leqq a''$, $b' \leqq y \leqq b''$ und $c' \leqq x \leqq c''$, $d' \leqq y \leqq d''$; es muß $a'' \leqq c'$ oder $c'' \leqq a'$ oder $b'' \leqq d'$ oder $d'' \leqq b'$ sein. Für $a'' \leqq c'$ ist $\mathfrak{M}(R, a'')$ Teil von $\mathfrak{M}(R, c')$, also $\mathfrak{N}(R, a'')$ orth. zu ihm — daher ist $P_{\mathfrak{N}(R, a'')} P_{\mathfrak{M}(R, c')} = 0$ und erst recht $E(A, \mathfrak{R}) \cdot E(A, \mathfrak{S}) = 0$. Ebenso schließt man in den drei anderen Fällen.

[71]) Es muß $a'' = c'$ oder $c'' = a'$ und gleichzeitig $b' = d'$, $d'' = b''$ sein, oder $b'' = d'$ oder $d'' = b'$ und gleichzeitig $a' = c'$, $a'' = c''$; man verifiziere an der Definition der $E(A, \mathfrak{R})$.

6. Es sei $0 \leqq \varrho < \sigma \leqq 1$, dann gibt es sicher Rechteck-Mengen $\mathfrak{U}$, die mit dem Einheitskreise genau den Bogen $2\pi\varrho, 2\pi\sigma$ gemein haben (und für die keine Randecke auf dem Einheitskreise liegt), und für alle diese ist (nach dem oben Gesagten) $E(A, \mathfrak{U})$ derselbe P. O. — er heiße $F(\varrho, \sigma)$. Es ist (für $0 \leqq \varrho < \sigma < \tau \leqq 1$) offenbar $F(\varrho, \sigma) + F(\varrho, \tau) = F(\varrho, \tau)$, also, wenn wir $F(\varrho) = F(0, \varrho)$ setzen, $F(\varrho, \sigma) = F(\sigma) - F(\varrho)$. Weiter muß, weil offenbar $F(0, 1) = 1$ ist, $F(1) = 1$ sein. Die $F(\varrho)$ sind P. O., da für $\varrho \leqq \sigma$ $F(\sigma) - F(\varrho)$ auch P. O. ist, ist $F(\varrho)$ Teil von $F(\sigma)$. $F(0)$ ist zunächst undefiniert, es sei $= 0$.

Man teile das Integrationsgebiet der Formel

$$(Af, f) = \iint (a + ib)\, d\, |E(A, a, b) f|^2$$

(ein großes Rechteck!) in Rechtecke $\mathfrak{R}_1, \ldots, \mathfrak{R}_k$ mit den folgenden Eigenschaften ein: Jedes derselben hat einen Durchmesser $\leqq \varepsilon$ ($\varepsilon > 0$), keines eine Randecke auf dem Einheitskreise; und zwar seien sie so angeordnet, daß mit dem Einheitskreise $\mathfrak{R}_1, \ldots, \mathfrak{R}_h$ ($h \leqq k$) Punkte gemein haben, und zwar bzw. die Bögen $2\pi\varrho_0, 2\pi\varrho_1; \ldots, 2\pi\varrho_{h-1}, 2\pi\varrho_h$ ($0 = \varrho_0 < \varrho_1 < \ldots < \varrho_{h-1} < \varrho_h = 1$). Aus jedem $\mathfrak{R}_1, \ldots, \mathfrak{R}_k$ wählen wir je einen Punkt $\xi_1, \ldots, \xi_k$ aus, dann ist

$$(Af, f) = \iint (a + ib)\, d\, |E(A, a, b) f|^2 = \sum_{\nu=1}^{k} \iint_{\mathfrak{R}_\nu} (a + ib)\, d\, |E(A, a, b) f|^2$$

beliebig wenig[72]) von

$$\sum_{\nu=1}^{k} \xi_\nu \iint_{\mathfrak{R}_\nu} d\, |E(A, a, b) f|^2 = \sum_{\nu=1}^{k} \xi_\nu \iint_{\mathfrak{R}_\nu} d\, (E(A, a, b) f, f) = \sum_{\nu=1}^{k} \xi_\nu (E(A, \mathfrak{R}_\nu) f, f)$$

$$= \sum_{\nu=1}^{h} \xi_\nu \cdot |F(\varrho_{\nu-1}, \varrho_\nu) f|^2 = \sum_{\nu=1}^{h} \xi_\nu \cdot (|F(\varrho_\nu) f|^2 - |F(\varrho_{\nu-1}) f|^2)$$

verschieden. Wir können natürlich $\xi_\nu = e^{2\pi i \cdot \varrho'_\nu}$, $\varrho_{\nu-1} < \varrho'_\nu < \varrho_\nu$ ($\nu = 1, \ldots, h$) wählen, dann ist unser Ausdruck

$$\sum_{\nu=1}^{h} e^{2\pi i \cdot \varrho'_\nu} (|F(\varrho_\nu) f|^2 - |F(\varrho_{\nu-1}) f|^2).$$

Dabei sind die $\varrho_0, \varrho_1, \ldots, \varrho_{h-1}, \varrho_h$ voneinander $\leqq \varepsilon$ entfernt.

Hieraus folgt aber

$$(Af, f) = \int_0^1 e^{2\pi i \cdot \varrho}\, d\, |F(\varrho) f|^2. \text{ }^{73})$$

---

[72]) Da $\iint_{\mathfrak{S}} d\, |E(A, a, b) f|^2$ für alle Rechtecke $\mathfrak{S} \geqq 0$ ist (nämlich $= |E(A, \mathfrak{S}) f|^2$), höchstens $\varepsilon \iint |E(A, a, b) f|^2 = \varepsilon \cdot |f|^2$.

[73]) Das Integral rechts konvergiert, z. B. weil wir $\varrho_0, \varrho_1, \ldots, \varrho_{h-1}, \varrho_h$ vorschreiben und die Einteilung $\mathfrak{R}_1, \ldots, \mathfrak{R}_k$ daran anpassen können.

Satz 12*. *Zu einem unitären Operator U gibt es eine Schar von P.O. $E(\lambda)$ $(0 \leqq \lambda \leqq 1)$ mit den folgenden Eigenschaften:*

α) Wenn $\lambda \leqq \mu$ ist, so ist $E(\lambda)$ Teil von $E(\mu)$; $E(0) = 0$, $E(1) = 1$.

β) Wenn $\lambda \geqq \lambda_0$, $\lambda \to \lambda_0$, so gilt für jedes f $E(\lambda)f \to E(\lambda_0)f$.

Dabei gilt für alle f, g

$$(Uf, g) = \int_0^1 e^{2\pi i \cdot \lambda} d(E(\lambda) f, g).$$

Beweis. Betrachten wir die vorhin konstruierten $F(\lambda)$. Sie haben die Eigenschaft α), aber β) eventuell nicht. Sei $0 \leqq \lambda_0 < 1$ und $\lambda_1 > \lambda_2 > \ldots > \lambda_0$, $\lambda_n \to \lambda$, dann bilden die $F(\lambda_1), F(\lambda_2), \ldots$ eine abnehmende Folge von P.O., nach Satz 19 gibt es also einen P.O. E, so daß stets $E(\lambda_n)f \to Ef$ ist; E hängt nur von λ_0 ab, denn zu zwei Folgen $\lambda_1, \lambda_2, \ldots$ muß dasselbe E gehören, weil geeignete Teilfolgen von ihnen zu einer ebensolchen Folge zusammengefaßt werden können — für die im obigen Sinne die Konvergenz auch stattfinden muß. Dieses E heiße $E(\lambda_0)$.

Da $F(\lambda_0)$ Teil aller $F(\lambda_1), F(\lambda_2), \ldots$ ist, ist es auch Teil von $E(\lambda_0)$; wenn $\lambda_0 < \mu_0$ ist, so können wir $\lambda_1 = \mu_0$ wählen, so daß $E(\lambda_0)$ Teil von $F(\mu_0) = F(\lambda_1)$ ist (weil es alle $F(\lambda_1), F(\lambda_2), \ldots$ sind). Also: $F(\lambda)$ Teil von $E(\lambda)$, für $\lambda < \mu$ $E(\lambda)$ Teil von $F(\mu)$, daher für $\lambda \leqq \mu$ $E(\lambda)$ Teil von $E(\mu)$.

Sei nun $\lambda > \lambda_0$, $\lambda \to \lambda_0$. $E(\lambda) - E(\lambda_0)$ ist Teil von $F(\lambda) - E(\lambda_0)$, also $|E(\lambda)f - E(\lambda_0)f| \leqq |F(\lambda)f - E(\lambda_0)f|$, und da letzteres (nach Definition von $E(\lambda_0)$) gegen 0 strebt, muß $E(\lambda)f \to E(\lambda_0)f$. Die $E(\lambda)$ erfüllen also β) und auch α) bis auf $E(0) = 0$, $E(1) = 1$. Wir erzwingen dies durch Definition — dann geht β) für $\lambda = 0$ eventuell verloren — das alte $E(0)$ heiße $\overline{E}$.

Für $F(\lambda)$ hatten wir die Relation

$$(Uf, f) = \int_0^1 e^{2\pi i \cdot \lambda} d\,|F(\lambda) f|^2.$$

Da $|F(\lambda)f|^2$, $|E(\lambda)f|^2$ monoton sind (in λ) und für $\lambda < \mu$ $|F(\lambda)f|^2 \leqq |E(\lambda)f|^2 \leqq |F(\mu)f|^2$ gilt, überlegt man sich leicht, daß sie auch für $E(\lambda)$ besteht. Dies ist aber die Schlußrelation des Satzes für $f = g$. Wenn man $\frac{f+g}{2}$ und $\frac{f-g}{2}$ einsetzt und subtrahiert, so erhält man ihren Realteil für beliebige f, g, ersetzt man f, g durch if, g, dann entsteht der Imaginärteil.

Es bleibt noch die Verletzung β) für $\lambda = 0$ zu beheben, dies gelingt, indem man für $0 < \lambda < 1$ $E(\lambda)$ durch $E(\lambda) - \overline{E}$ ersetzt — dann bleibt alles andere beim alten und dieser Punkt kommt in Ordnung.

7. Der Satz 12* war die Hauptschwierigkeit, nun folgen einige leichte Eindeutigkeitssätze.

Satz 13*. *Wenn die P. O. $E(\lambda)$ den Bedingungen α), β) von Satz 12* entsprechend gewählt werden, so gibt es genau einen unitären Operator U, der zu ihnen im Verhältnis von Satz 12* steht.*

Beweis. Daß es höchstens ein solches U gibt, ist klar: Satz 12* legt alle (Uf, g), also alle Uf fest. Bleibt die Existenz zu beweisen.

Durch dieselben Überlegungen, wie sie beim Beweise von Satz 36 angestellt wurden, gelingt (wegen $|e^{2\pi i \cdot \lambda}| = 1$) die Abschätzung

$$\left|\int_0^1 e^{2\pi i \cdot \lambda}\, d(E(\lambda) f, g)\right| \leqq |f|\,|g|.$$

Nennen wir das Integral links $L(g)$ (f fest, g variabel), so ist

$$L(a_1 g_1 + \ldots + a_n g_n) = \bar{a}_1 L(g_1) + \ldots + \bar{a}_n L(g_n),$$
$$|L(g)| \leqq |f|\,|g| = C \cdot |g|,$$

der F. Rieszsche Satz ist anwendbar (vgl. Anm. [52])): Es gibt genau ein f^* mit $L(g) = (f^*, g)$. Wir definieren U durch $Uf = f^*$, es ist offenbar ein überall sinnvoller lin. O. — wir müssen nur noch die Unitarität von U beweisen.

Zunächst ist $|(Uf, g)| \leqq |f|\,|g|$, also U beschränkt; wir bilden U^*, und haben:

$$(Uf, g) = \int_0^1 e^{2\pi i \lambda}\, d(E(\lambda) f, g), \qquad (U^* f, g) = \int_0^1 e^{-2\pi i \lambda}\, d(E(\lambda) f, g)$$

(die erste Gleichung ist Definition, die zweite entsteht durch Vertauschen von f, g und Nehmen der komplex-Konjugierten). Genau wie am Ende von § 4 berechnet man auf Grund dieser Formeln $(Uf, Ug) = (f, g)$, $(U^* f, U^* g) = (f, g)$. Daraus folgt $(U^* U f, g) = (U U^* f, g) = (f, g)$ für alle f, g, also $U^* U = U U^* = 1$ — also die Unitarität.

Satz 14*. *Zu einem gegebenen unitären Operator U kann nur eine Schar $E(\lambda)$ von P. O. im Verhältnisse des Satzes 12* stehen.*

Beweis. Die Formel

$$(U^n f, g) = \int_0^1 e^{2n\pi i \lambda}\, d(E(\lambda) f, g)$$

ist für $n = 0$ trivial, für $n = 1$ Definition, für $n = -1$ Folge von $U^{-1} = U^*$. Wenn sie für m, n bewiesen ist, so gilt sie auch für $m - n$: denn mit der Methode vom Schlusse von § 4 ergibt sich

$$(U^m f, U^n g) = \int_0^1 e^{2(m-n)\pi i \lambda}\, d(E(\lambda) f, g),$$

und es ist

$$(U^m f, U^n g) = ((U^n)^* U^m f, g) = (U^{*n} U^m f, g) = (U^{-n} U^m f, g) = (U^{m-n} f, g).$$

Daher gilt sie für alle $n = 0, \pm 1, \pm 2, \ldots$.

Wenn $F(\lambda)$ eine zweite Schar von P.O. mit denselben Eigenschaften ist, so gilt die obige Gleichung auch mit $F(\lambda)$ an Stelle von $E(\lambda)$. Also gilt

$$\int_0^1 p(\lambda)\, d((E(\lambda) f, g) - (F(\lambda) f, g)) = 0$$

zunächst für alle $p(\lambda) = e^{2n\pi i\lambda}$, und dann auch für ihre Linearaggregate, die trigonometrischen Polynome. Aus Stetigkeitsgründen gilt es aber auch für alle stetigen Funktionen $p(\lambda)$ mit $p(0) = p(1)$. Wegen der Bedingung β) in Satz 12* gilt sie sogar noch für solche $p(\lambda)$, die endlich viele Sprünge aufweisen — vorausgesetzt, daß die Unstetigkeit stets rechts von der Sprungstelle liegt. Wir setzen nun

$$p(\lambda) = \begin{cases} 1, & \text{für} \quad 0 < \lambda \leqq \lambda_0, \\ 0, & \text{sonst} \end{cases},$$

$(0 \leqq \lambda_0 < 1)$, dann haben wir

$$\int_0^{\lambda_0} d((E(\lambda) f, g) - (F(\lambda) f, g)) = (E(\lambda_0) f, g) - (F(\lambda_0) f, g) = 0;$$

da dies für alle f, g gilt, ist somit $E(\lambda_0) = F(\lambda_0)$. (Dabei muß $\lambda_0 \neq 1$ sein, für $\lambda_0 = 1$ ist es aber wegen α) in Satz 12* der Fall.)

8. Wir haben durch Satz 12*, 13*, 14* gezeigt: die von Satz 12* gegebene Zuordnung eines unitären Operators U zu einer Schar von P.O. $E(\lambda)$ mit den dortigen Eigenschaften α), β) ist ein-eindeutig, und alle U und alle $E(\lambda)$ werden dadurch erschöpft. Zur Z. d. E. fehlt diesen $E(\lambda)$ noch die Eigenschaft $E(\lambda) f \to f$ für $\lambda < 1, \lambda \to 1$ (vgl. Def. 17), wir haben daher alles in Kap. IX Angekündigte bewiesen, wenn wir noch zeigen, daß gerade diese Eigenschaft damit gleichbedeutend ist, daß $\varphi = U\varphi$ nur für $\varphi = 0$ stattfindet.

Sei zuerst $\varphi = U\varphi$, $\varphi \neq 0$. Dann ist

$$\begin{aligned} \Re((\varphi, \varphi) - (U\varphi, \varphi)) &= \Re \int_0^1 (1 - e^{2\pi i \cdot \lambda})\, d(E(\lambda)\varphi, \varphi) \\ &= \int_0^1 2 \sin^2 \pi\lambda\, d|E(\lambda)\varphi|^2. \end{aligned}$$

Dies muß verschwinden, und ist

$$\begin{aligned} &\geqq \int_\varepsilon^{1-\varepsilon} 2\sin^2\pi\lambda\, d|E(\lambda)\varphi|^2 \geqq 2\sin^2\pi\varepsilon \int_\varepsilon^{1-\varepsilon} d|E(\lambda)\varphi|^2 \\ &= 2\sin^2\pi\varepsilon\,(|E(1-\varepsilon)\varphi|^2 - |E(\varepsilon)\varphi|^2) \end{aligned}$$

(für alle $\varepsilon > 0$). Für $\varepsilon \leqq \frac{1}{2}$ ist die rechte Seite gewiß $\geqq 0$, also $= 0$:

$$|E(1-\varepsilon)\varphi| = |E(\varepsilon)\varphi|.$$

Da die rechte Seite mit $\varepsilon \to 0$ gegen 0 strebt, gilt dasselbe von der linken: $E(1-\varepsilon)\varphi \to 0 \neq \varphi$.

Sei umgekehrt für $\lambda < 1$, $\lambda \to 1$ nicht $E(\lambda)f \to f$. Nach Satz 19 existiert ein P.O. E, so daß für alle g $E(\lambda)g \to Eg$, es ist also $Ef \neq f$. Für $f' = f - Ef$ haben wir daher: $f' \neq 0$, $E(\lambda)f' \to Ef' = 0$. Nun ist $|E(\lambda)f'|$ nie negativ, mit λ monoton nicht fallend — da es für $\lambda \to 1$ gegen 0 strebt, muß es stets $= 0$ sein. Also ist $E(\lambda)f' = 0$ für alle $\lambda < 1$, und natürlich $E(1)f' = f'$. Hieraus folgt

$$(Uf', g) = \int_0^1 e^{2\pi i \cdot \lambda} d(E(\lambda)f', g) = (f', g),$$

dies gilt für alle g, also ist $Uf' = f'$.

Damit ist alles bewiesen.

Anhang III. Operatoren und Matrizen.

1. Es ist wohl erwünscht, unsere Resultate über H. O. in die übliche Terminologie der Matrizen zu übersetzen. Es wird sich zeigen, daß im Unbeschränkten die der Operatoren bedeutend praktischer ist. (Im Beschränkten kommt beides, nach Hilberts Resultaten, aufs selbe heraus.) Im übrigen werden wir erkennen, daß die abg. lin. H. O. und die Hermiteschen Matrizen mit lauter endlichen (aber nicht notwendig beschränkten!) Zeilen-Absolutwertquadratsummen — die sogenannten „quadrierbaren Matrizen“ — einander zugeordnet werden können.

Die genauere Theorie dieser Matrizen soll in der in Anm.[22]) angekündigten Arbeit ausgeführt werden, hier soll nur soviel gegeben werden, als zur Orientierung nötig ist, und was wir für die Zwecke von Anhang IV brauchen.

Eine Matrix $A = \{a_{\mu\nu}\}$ $(\mu, \nu = 1, 2, \ldots)$ heiße Hermitesch quadrierbar (kurz: H. quadr.), wenn $a_{\mu\nu} = \overline{a_{\nu\mu}}$ gilt, und alle $\sum_{\nu=1}^{\infty} |a_{\mu\nu}|^2$ $(\mu = 1, 2, \ldots)$ endlich sind[74]). Wenn $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System ist, so bilden wir einen Operator R, der für die φ_μ (und nur diese) Sinn hat, und zwar $R\varphi_\mu = \sum_{\nu=1}^{\infty} a_{\mu\nu}\varphi_\nu$ (die Reihe konvergiert). Die von den $\varphi_1, \varphi_2, \ldots$ aufgespannte abg. lin. M. ist $\mathfrak{H}$, und es ist

$$(\varphi_\mu, R\varphi_\nu) = \overline{a_{\nu\mu}} = a_{\mu\nu} = (R\varphi_\mu, \varphi_\nu),$$

[74]) Dann sind auch alle $\sum_{\varrho=1}^{\infty} a_{\mu\varrho}\overline{a_{\nu\varrho}}$ absolut konvergent, d. h. die Matrix A^2 kann gebildet werden.

d. h. R ist ein H. O. Wir nennen R den elementaren H. O. von $A, \{\varphi_1, \varphi_2, \ldots\}$; und $\hat{R}, \tilde{R}$ den lin. bzw. abg. lin. H. O. von $A, \{\varphi_1, \varphi_2, \ldots\}$. — $\hat{R}f$ hat offenbar für die $\sum_{\mu=1}^{n} x_\mu \varphi_\mu$ Sinn (und nur für diese), und es ist

$$\hat{R}f = \sum_{\mu=1}^{n} x_\mu \left(\sum_{\mu=1}^{\infty} a_{\mu\nu} \varphi_\nu\right) = \sum_{\nu=1}^{\infty} \left[\sum_{\mu=1}^{n} a_{\mu\nu} x_\mu\right] \varphi_\nu.$$

$\tilde{R}$ ist nicht so leicht direkt zu charakterisieren, wohl aber seine Erweiterungselemente. Damit f eines sei, und zwar das f^* ihm zugeordnet, muß

$$(f^*, \varphi_\nu) = (f, R\varphi_\nu) = \sum_{\mu=1}^{\infty} (f, \varphi_\mu)(\varphi_\mu, R\varphi_\nu) = \sum_{\mu=1}^{\infty} a_{\mu\nu} (f, \varphi_\mu) \qquad (\mu = 1, 2, \ldots)$$

sein (R und $\tilde{R}$ haben ja dieselben Erweiterungselemente). Wir schreiben

$$(f, \varphi_\mu) = x_\mu \qquad (\mu = 1, 2, \ldots), \qquad f = \sum_{\mu=1}^{\infty} x_\mu \varphi_\mu,$$

dann existiert ein f^* mit $(f^*, \varphi_\nu) = \sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu$ offenbar dann und nur dann, wenn $\sum_{\nu=1}^{\infty} \left|\sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu\right|^2$ endlich ist, und zwar ist es dann gleich $\sum_{\nu=1}^{\infty} \left(\sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu\right) \varphi_\nu$. Wir haben also: $f = \sum_{\mu=1}^{\infty} x_\mu \varphi_\mu \left(\sum_{\mu=1}^{\infty} |x_\mu|^2 \text{ endlich}\right)$ ist dann und nur dann Erweiterungselement, wenn für die

$$y_\nu = \sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu \text{ [75])}$$

$\sum_{\mu=1}^{\infty} |y_\mu|^2$ endlich ist, und zwar ist dann $f^* = \sum_{\mu=1}^{\infty} y_\mu \varphi_\mu$. (Wenn sogar $\tilde{R}f$ Sinn hat, ist es erst recht dieses.)

Weiter ist

$$(f^*, f) = \sum_{\nu=1}^{\infty} y_\nu \bar{x}_\nu = \sum_{\nu=1}^{\infty} \left(\sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu\right) \bar{x}_\nu = \sum_{\nu=1}^{\infty} \left(\sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu \bar{x}_\nu\right),$$

$$(f, f^*) = \sum_{\mu=1}^{\infty} x_\mu \bar{y}_\mu = \sum_{\mu=1}^{\infty} x_\mu \left(\sum_{\nu=1}^{\infty} \bar{a}_{\nu\mu} \bar{x}_\nu\right) = \sum_{\mu=1}^{\infty} \left(\sum_{\nu=1}^{\infty} a_{\mu\nu} x_\mu \bar{x}_\nu\right).$$

Die 0-Klasse ist also durch die Vertauschbarkeit von $\sum_{\mu=1}^{\infty}, \sum_{\nu=1}^{\infty}$ gekennzeichnet!

2. Wir zeigen nun, daß zu jedem abg. lin. H. O. R eine H. quadr. Matrix $A = \{a_{\mu\nu}\}$ und ein vollst. norm. orth. System $\varphi_1, \varphi_2, \ldots$ existiert, so daß R der abg. lin. H. O. von $A, \{\varphi_1, \varphi_2, \ldots\}$ ist. Es genügt $\varphi_1, \varphi_2, \ldots$ anzugeben, A ist ja durch $a_{\mu\nu} = (R\varphi_\mu, \varphi_\nu)$ festgelegt, und wenn die $R\varphi_1, R\varphi_2, \ldots$ alle Sinn haben, ist A H. quadr.: $a_{\mu\nu} = \overline{a_{\mu\nu}}$ ist klar, und $\sum_{\nu=1}^{\infty} |a_{\mu\nu}|^2 = \sum_{\nu=1}^{\infty} |(R\varphi_\mu, \varphi_\nu)|^2 = |R\varphi_\mu|^2$ ist endlich.

---

[75]) $\sum_{\mu=1}^{\infty} |a_{\mu\nu}|^2$, $\sum_{\mu=1}^{\infty} |x_\mu|^2$ sind endlich, also konvergieren die $\sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu$ jedenfalls absolut!

Sei S der elementare Operator von $A, \{\varphi_1, \varphi_2, \ldots\}$, dann ist

$$S\varphi_\mu = \sum_{\nu=1}^{\infty} a_{\mu\nu}\varphi_\nu = \sum_{\nu=1}^{\infty} (R\varphi_\mu, \varphi_\nu)\cdot\varphi_\nu = R\varphi_\mu;$$

d. h. S ist das auf die $\varphi_1, \varphi_2, \ldots$ eingeschränkte R. Das, was wir durch geeignete Wahl der $\varphi_1, \varphi_2, \ldots$ erreichen sollen, ist also $\tilde{S} = R$, d. h.: Daß zu jedem f mit sinnvollem Rf eine Folge von Linearaggregaten je endlichvieler $\varphi_1, \varphi_2, \ldots$ existiere, $f_1, f_2, \ldots$, so daß $f_n \to f$, $Rf_n \to Rf$ (natürlich müssen alle $R\varphi_n$ sinnvoll sein). Dies ist erreicht, wenn wir eine Folge $g_1, g_2, \ldots$ mit lauter sinnvollen $Rg_1, Rg_2, \ldots$ haben, derart, daß zu jedem f mit sinnvollem Rf eine Teilfolge $g_{N_1}, g_{N_2}, \ldots$ mit $g_{N_n} \to f$, $Rg_{N_n} \to Rf$ existiert (diese ist dann von selbst überall dicht): Dann genügt es, $f_1, f_2, \ldots$ nach Satz 8 zu $\varphi_1, \varphi_2, \ldots$ zu „orthogonalisieren".

Eine solche Folge $g_1, g_2, \ldots$ finden wir so: Sei $h_1, h_2, \ldots$ überall dicht, aber sonst beliebig. Zu jedem Zahlentripel m, n, k $(= 1, 2, \ldots)$ zu dem es ein f mit sinnvollem Rf und $|h_m - f| \leqq \frac{1}{k}$, $|h_n - Rf| \leqq \frac{1}{k}$ gibt, ordnen wir ein solches f zu: $f_{m,n,k}$. Diese $f_{m,n,k}$ leisten (als Folge geschrieben) alles Gewünschte. Denn sei f, Rf und $\varepsilon > 0$ gegeben. Wir wählen ein k mit $\frac{2}{k} \leqq \varepsilon$ und h_m und h_n mit $|h_n - f| \leqq \frac{1}{k}$, $|h_m - Rf| \leqq \frac{1}{k}$. Dann ist $f_{m,n,k}$ gewiß definiert, und $|h_m - f_{m,n,k}| \leqq \frac{1}{k}$, $|h_n - Rf_{m,n,k}| \leqq \frac{1}{k}$, also $|f - f_{m,n,k}| \leqq \varepsilon$, $|Rf - Rf_{m,n,k}| \leqq \varepsilon$. Da $\varepsilon > 0$ beliebig ist, sind wir am Ziele.

3. Zu gegebenem R kann es auch mehrere $A, \{\varphi_1, \varphi_2, \ldots\}$ geben, deren abg. lin. H. O. dieses R ist. Seien etwa $A = \{a_{\mu\nu}\}, \{\varphi_1, \varphi_2, \ldots\}$ und $B = \{b_{\mu\nu}\}, \{\psi_1, \psi_2, \ldots\}$ derart. (Sowohl A wie B sollen H. quadr. sein!) Die beiden vollst. norm. orth. Systeme $\varphi_1, \varphi_2, \ldots, \psi_1, \psi_2, \ldots$ bestimmen eine unitäre unendliche Matrix $\{u_{\mu\nu}\}$ mit $u_{\mu\nu} = (\varphi_\mu, \psi_\nu)$. In der Tat ist:

$$(\mathrm{U}_1)\quad \begin{aligned} \sum_{\varrho=1}^{\infty} u_{\mu\varrho}\overline{u_{\nu\varrho}} &= \sum_{\varrho=1}^{\infty} (\varphi_\mu, \psi_\varrho)(\psi_\varrho, \varphi_\nu) = (\varphi_\mu, \varphi_\nu) = \begin{cases} 1, & \text{für } \mu = \nu, \\ 0, & \text{für } \mu \neq \nu, \end{cases} \\ \sum_{\varrho=1}^{\infty} u_{\varrho\mu}\overline{u_{\varrho\nu}} &= \sum_{\varrho=1}^{\infty} (\psi_\nu, \varphi_\varrho)(\varphi_\varrho, \psi_\mu) = (\psi_\nu, \psi_\mu) = \begin{cases} 1, & \text{für } \mu = \nu, \\ 0, & \text{für } \mu \neq \nu. \end{cases} \end{aligned}$$

Und weiter:

$$(\mathrm{U}_2)\quad \varphi_\mu = \sum_{\nu=1}^{\infty} (\varphi_\mu, \psi_\nu)\cdot\psi_\nu = \sum_{\nu=1}^{\infty} u_{\mu\nu}\cdot\psi_\nu,\quad \psi_\mu = \sum_{\nu=1}^{\infty} (\psi_\mu, \varphi_\nu)\cdot\varphi_\nu = \sum_{\nu=1}^{\infty} \overline{u_{\nu\mu}}\cdot\varphi_\nu.$$

Auch für $a_{\mu\nu}$, $b_{\mu\nu}$ gelten die gewohnten Transformationsformeln:

$$\begin{aligned} a_{\mu\nu} = (\varphi_\mu, R\varphi_\nu) &= \sum_{\varrho=1}^{\infty} (\varphi_\mu, \psi_\varrho)(\psi_\varrho, R\varphi_\nu) = \sum_{\varrho=1}^{\infty} (\varphi_\mu, \psi_\varrho)(R\psi_\varrho, \varphi_\nu) \\ &= \sum_{\varrho=1}^{\infty} (\varphi_\mu, \psi_\varrho)\Big[\sum_{\sigma=1}^{\infty} (R\psi_\varrho, \psi_\sigma)(\psi_\sigma, \varphi_\nu)\Big] = \sum_{\varrho=1}^{\infty} \Big(\sum_{\sigma=1}^{\infty} b_{\varrho\sigma}\, u_{\mu\varrho}\overline{u_{\nu\sigma}}\Big). \end{aligned}$$

Wenn wir μ, ν (und ϱ, σ) vertauschen und die komplex-Konjugierten nehmen, so wird (wegen $a_{\mu\nu} = \overline{a_{\nu\mu}}$, $b_{\varrho\sigma} = \overline{b_{\sigma\varrho}}$)

$$(\mathrm{U}_3) \qquad a_{\mu\nu} = \sum_{\varrho=1}^{\infty}\Big(\sum_{\sigma=1}^{\infty} b_{\varrho\sigma} u_{\mu\varrho} \overline{u_{\nu\sigma}}\Big) = \sum_{\sigma=1}^{\infty}\Big(\sum_{\varrho=1}^{\infty} b_{\varrho\sigma} u_{\mu\varrho} \overline{u_{\nu\sigma}}\Big).$$

Ebenso erhalten wir die Umkehrungen:

$$(\mathrm{U}_4) \qquad b_{\mu\nu} = \sum_{\varrho=1}^{\infty}\Big(\sum_{\sigma=1}^{\infty} a_{\varrho\sigma} \overline{u_{\varrho\mu}}\, u_{\sigma\nu}\Big) = \sum_{\sigma=1}^{\infty}\Big(\sum_{\varrho=1}^{\infty} b_{\varrho\sigma} \overline{u_{\varrho\mu}}\, u_{\sigma\nu}\Big).$$

Bis hierher gilt also die gewöhnliche unitäre Transformationstheorie der Hermiteschen Matrizen, aber nicht weiter! Denn wenn wir eine H. quadr. Matrix $A = \{a_{\mu\nu}\}$ und ein vollst. norm. orth. System $\varphi_1, \varphi_2, \ldots$ haben, und irgendeine unitäre Transformationsmatrix $\{u_{\mu\nu}\}$ (nach (U_1)) darauf anzuwenden versuchen (nach (U_2), (U_3), (U_4)), so braucht zunächst nichts zu konvergieren. Aber auch wenn die in (U_3), (U_4) angegebenen Konvergenzen alle stattfinden, so kann sich der abg. lin. H. O. von $A_1\{\varphi_1, \varphi_2, \ldots\}$ bei der Transformation ändern. Wir gehen auf diese Verhältnisse hier nicht näher ein, da sie an anderem Orte (vgl. Anm. [22]) genau diskutiert werden sollen.

4. Es sei noch erwähnt, wie sich die abg. lin. M. $\mathfrak{E}$, $\mathfrak{F}$ die zum abg. lin. H. O. $\tilde{R}$ von $A, \{\varphi_1, \varphi_2, \ldots\}$ gehören (R sei sein elementarer H. O.) bestimmen lassen (vgl. Satz 22). $\mathfrak{E}$ bzw. $\mathfrak{F}$ ist die Menge aller $(\tilde{R} \pm i1)f$. Aus der abg. von $\tilde{R}$ und den in Satz 22 auseinandergesetzten Eigenschaften von $\tilde{R} \pm i1$ folgert man leicht, daß dies die Menge der Häufungspunkte von $(\hat{R} \pm i1)f$ ist[76]. Dies aber ist die von den $(R \pm i1)f$, d. h. von den $(R \pm i1)\varphi_\mu$, aufgespannte lin. M.; also ist $\mathfrak{E}$ bzw. $\mathfrak{F}$ die von den $(R \pm i1)\varphi_\mu$ aufgespannte abg. lin. M., d. h. von den $\sum_{\nu=1}^{\infty} a_{\mu\nu}\varphi_\nu \pm i\varphi_\mu$.

Anhang IV. Der Operator $\overline{R}$.

1. $\overline{R}$ und seine Cayleysche Transformierte $\overline{U}$ wurden in Kap. X (Satz 37) folgendermaßen definiert: Sei $\varphi_0, \varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System, dann ist

$$\overline{U}\Big(\sum_{\nu=0}^{\infty} x_\nu \varphi_\nu\Big) = \sum_{\nu=0}^{\infty} x_\nu \varphi_{\nu+1} \qquad \Big(\sum_{\nu=1}^{\infty} |x_\nu|^2 \text{ endlich}\Big);$$

d. h. $\overline{R}f$ hat Sinn für alle $f = \varphi - \overline{U}\varphi = x_0\varphi_0 + (x_1 - x_0)\varphi_1 + (x_2 - x_1)\varphi_2 + \ldots$, und dann ist $\overline{R}f = i(\varphi + \overline{U}\varphi) = i x_0 \varphi_0 + i(x_0 + x_1)\varphi_1 + i(x_1 + x_2)\varphi_2 + \ldots$. Oder auch: $\overline{R}\Big(\sum_{\nu=0}^{\infty} y_\nu \varphi_\nu\Big)$ hat Sinn, wenn $|y_0|^2 + |y_0 + y_1|^2 + |y_0 + y_1 + y_2|^2 + \ldots$

[76]) Denn wenn eine Folge $(\tilde{R} \pm i1)f$ konvergiert, so konvergiert auch $U^{\pm 1}(\tilde{R} \pm i1)f = (\tilde{R} \mp i1)f$, also f und $\tilde{R}f$ — und natürlich auch umgekehrt: Wenn $f, \tilde{R}f$ konvergieren, konvergiert $(\tilde{R} \pm i1)f$.

konvergiert (es muß ja $x_0 = y_0$, $x_1 = y_0 + y_1$, $x_2 = y_0 + y_1 + y_2, \ldots$ sein), und zwar ist es dann gleich $i y_0 \varphi_0 + (2 i y_0 + i y_1) \varphi_1 + (2 i y_0 + 2 i y_1 + i y_2) \varphi_2 + \ldots$. Wir haben hier eine Art Matrix von $\bar{R}$ gewonnen:

$$a_{\mu\nu} = \begin{cases} 0, & \text{für } \mu > \nu, \\ i, & \text{für } \mu = \nu, \\ 2i, & \text{für } \mu < \nu. \end{cases}$$

Freilich ist sie es nicht im Sinne von Anhang III: Sie ist weder H. noch quadr., was kein Wunder ist, da die $\bar{R}\varphi_\mu$, $\mu = 0, 1, 2, \ldots$, alle sinnlos sind.

Um eine richtige Matrix von $\bar{R}$ angeben zu können, müssen wir vor allem ein vollst. norm. orth. System $\psi_1, \psi_2, \ldots$ finden, für welches alle $\bar{R}\psi_\mu$, $\mu = 1, 2, \ldots$, Sinn haben.

Wir setzen:

$$\omega_{m,k} = \varphi_{m\cdot 2^{k+1}} + \varphi_{m\cdot 2^{k+1}+1} + \cdots + \varphi_{(2m+1)\cdot 2^k - 1} - \varphi_{(2m+1)\cdot 2^k}$$
$$- \varphi_{(2m+1)\cdot 2^k+1} - \cdots - \varphi_{(m+1)\cdot 2^{k+1}-1} \quad (m, k = 0, 1, 2, \ldots).$$

Nach unserem Kriterium ist $\bar{R}\left(\sum_{\nu=0}^{\infty} x_\nu \varphi_\nu\right)$ gewiß sinnvoll, wenn nur endlich viele $x_\nu \neq 0$ sind und $\sum_{\nu=1}^{\infty} x_\nu = 0$ ist — dies ist aber hier der Fall. $(\omega_{m,k}, \omega_{n,l}) = 0$ gilt sicher, wenn $\omega_{m,k}, \omega_{n,l}$ kein gemeinsames φ_ν-Glied haben, und auch wenn alle φ_ν des einen im anderen vorkommen, und zwar alle mit dem Koeffizienten $+1$, oder alle mit -1. Dies ist aber immer der Fall, außer für $m = n$, $k = l$, dagegen ist $|\omega_{m,k}|^2$ offenbar $= 2^{k+1}$. Die $\psi_{m,k} = 2^{-\frac{1}{2}(k+1)} \omega_{m,k}$ bilden also ein norm. orth. System, und alle $\bar{R}\psi_{m,k}$ sind sinnvoll. Aber dieses System ist auch vollst., denn sei $f = \sum_{\nu=0}^{\infty} x_\nu \varphi_\nu$ zu allen $\psi_{m,k}$, d. h. zu allen $\omega_{m,k}$ orth. Das bedeutet:

$$x_{m\cdot 2^{k+1}} + x_{m\cdot 2^{k+1}+1} + \cdots + x_{(2m+1)\cdot 2^k - 1} - x_{(2m+1)\cdot 2^k} - x_{(2m+1)\cdot 2^k - 1} - \cdots$$
$$- x_{(m+1)\cdot 2^{k+1}-1} = 0.$$

Also: $x_0 - x_1 = 0$, $x_2 - x_3 = 0$, $x_5 - x_4 = 0$, $x_7 - x_6 = 0, \ldots$, d. h. $x_0 = x_1$, $x_2 = x_3$, $x_4 = x_5$, $x_6 = x_7, \ldots$ Weiter: $x_0 + x_1 - x_2 - x_3 = 0$, $x_4 + x_5 - x_6 - x_7 = 0, \ldots$, d. h. $2x_0 = 2x_2$, $2x_4 = 2x_6, \ldots$, d. h. $x_0 = x_1 = x_2 = x_3$, $x_4 = x_5 = x_6 = x_7, \ldots$ Weiter: $x_0 + x_1 + x_2 + x_3 - x_4 - x_5 - x_6 - x_7 = 0$, d. h. $4x_0 = 4x_4, \ldots$, d. h. $x_0 = x_1 = \ldots = x_7, \ldots$. Somit gilt überhaupt $x_0 = x_1 = x_2 = \ldots$, und $\sum_{\nu=0}^{\infty} |x_\nu|^2$ kann nur endlich sein, wenn alle $= 0$ sind: also $f = 0$. Damit ist die Vollst. bewiesen.

2. Wenn wir nun die Matrix $\{b_{mk/nl}\}$ durch

$$b_{mk/nl} = (\psi_{m,k}, \bar{R}\psi_{n,l}) = 2^{-\frac{1}{2}(k+l)-1} (\omega_{m,k}, \bar{R}\omega_{n,l})$$

definieren, so ist sie jedenfalls H. quadr. und ihr elementarer H. O., S, ist das auf die $\psi_{m,k}$ eingeschränkte $\bar{R}$. Also ist $\bar{R}$ Forts. von S, und (weil es abg. lin. ist) von $\tilde{S}$. Wenn $\tilde{S}$ max. ist, so muß $\bar{R}=\tilde{S}$ sein, d. h. der abg. lin. H. O. von $\{b_{mk/nl}\}$, $\{\psi_{mk}\}$.

Wir beweisen dies, indem wir zeigen: das $\mathfrak{E}$ von $\tilde{S}$ ist $=\mathfrak{H}$, d. h. (vgl § 4 von Anhang III) die $\bar{R}\psi_{m,k}+i\psi_{m,k}$, oder auch die $\bar{R}\omega_{m,k}+i\omega_{m,k}$, spannen die abg. lin. M. $\mathfrak{H}$ auf. Oder: wenn $f=\sum_{\nu=0}^{\infty} x_\nu \varphi_\nu$ zu allen $\bar{R}\omega_{m,k}+i\omega_{m,k}$ orth. ist, so ist $f=0$. Nach § 1 können wir ja ausrechnen:

$$\begin{aligned}\bar{R}\omega_{m,k} = {} & i\varphi_{m\cdot 2^{k+1}}+3i\varphi_{m\cdot 2^{k+1}+1}+\dots+(2^{k+1}-1)\,i\varphi_{(2m+1)\cdot 2^k-1}\\ & +(2^{k+1}-1)\,i\varphi_{(2m+1)\cdot 2^k}+(2^{k+1}-3)\,i\varphi_{(2m+1)\cdot 2^k+1}+\dots\\ & +3i\varphi_{(m+1)\cdot 2^{k+1}-2}+i\varphi_{(m+1)\cdot 2^{k+1}-1},\end{aligned}$$

$$\begin{aligned}\bar{R}\omega_{m,k}+i\omega_{m,k} = {} & 2i\varphi_{m\cdot 2^{k+1}}+4i\varphi_{m\cdot 2^{k+1}+1}+\dots+2^{k+1}i\varphi_{(2m+1)\cdot 2^k-1}\\ & +(2^{k+1}-2)\,i\varphi_{(2m+1)\cdot 2^k}+(2^{k+1}-4)\,i\varphi_{(2m+1)\cdot 2^k+1}+\dots\\ & +2i\varphi_{(m+1)\cdot 2^{k+1}-2}.\end{aligned}$$

Daß f zu diesen orth. ist, bedeutet

$$2i\,x_{m\cdot 2^{k+1}}+4i\,x_{m\cdot 2^{k+1}+1}+\dots+2^{k+1}i\,x_{(2m+1)\cdot 2^k-1}+\dots+2i\,x_{(m+1)\cdot 2^{k+1}-2}=0,$$

d. h.

$$\begin{gathered}x_0=0,\ x_2=0,\ x_4=0,\dots;\ x_0+2x_1+x_2=0,\ x_4+2x_5+x_6=0,\dots;\\ x_0+2x_1+3x_2+4x_3+3x_4+2x_5+x_6=0,\dots;\ \dots,\end{gathered}$$

woraus $x_0=x_1=x_2=\dots 0$, $f=0$ folgt.

Es bleibt noch übrig, die $b_{m,k/n,l}=2^{-\frac{1}{2}(k+l)-1}(\omega_{m,k},\bar{R}\omega_{n,l})$ zu berechnen, was ohne weiteres geht, da wir $\omega_{m,k}$, $\bar{R}\omega_{n,l}$ mit den φ_ν ausgedrückt haben. Für $m\neq n$, $k=l$ treten keine gemeinsamen φ_ν auf, es kommt also 0 heraus, für $m=n$, $k=l$ überzeugt man sich direkt von demselben. Für $k>l$ und n außerhalb des Intervalles $m\cdot 2^{k-l},\dots,(m+1)\cdot 2^{k-l}-1$ fehlen wieder die gemeinsamen φ_ν, also 0; liegt n im genannten Intervalle, so ist es $=m\cdot 2^{k-l}+\varrho$ oder $=(m+1)\cdot 2^{k-l}-\varrho-1$ ($\varrho=0,1,\dots,2^{k-l-1}-1$), und dann berechnet man:

$$(\omega_{m,k},\bar{R}\omega_{n,l})=\begin{cases}\displaystyle\sum_{\varrho\cdot 2^{l+1}}^{(2\varrho+1)\cdot 2^l-1}{}_{\alpha}-i(2\alpha+1)-\sum_{(2\varrho+1)\cdot 2^l}^{(\varrho+1)\cdot 2^{l+1}-1}{}_{\alpha}-i(2\alpha+1)\\ \qquad =-i\cdot 2^l\cdot(-2\cdot 2^l)=i\cdot 2^{2l+1},\ \text{bzw.}\\ \displaystyle\sum_{(2\varrho+1)\cdot 2^l}^{(\varrho+1)\cdot 2^l-1}{}_{\alpha}-i(2\alpha+1)-\sum_{\varrho\cdot 2^{l+1}}^{(2\varrho+1)\cdot 2^l-1}{}_{\alpha}-i(2\alpha+1)\\ \qquad =-i\cdot 2^l\cdot 2\cdot 2^l=-i\cdot 2^{2l+1}.\end{cases}$$

Für $k < l$ muß es (weil $b_{mk/nl}$ Hermitesch ist) nach Vertauschen vom m, n (und k, l) entsprechend $0, -i\cdot 2^{2k+1}, i\cdot 2^{2k+1}$ sein. Und den $0, \pm i\cdot 2^{2k+1}$, $\pm i\cdot 2^{2l+1}$ entspricht bei $b_{mk/nl}$ $0, \pm i\cdot 2^{\frac{1}{2}(3k-l)}, \pm i\cdot 2^{\frac{1}{2}(3l-k)}$.

Wir führen nun die Indizierung m, k $(=0, 1, 2, \ldots)$ auf eine mit p $(=1, 2, \ldots)$ zurück, durch $p = (2m+1)\cdot 2^k$. Dann wird, wie man sich leicht überlegt:

Satz 1**. *Im oben beschriebenen vollst. norm. orth. System $\psi_1, \psi_2, \ldots$ gibt es eine H. quadr. Matrix $B = \{b_{pq}\}$, so daß $\bar{R}$ der abg. lin. H. O. von B, $\{\psi_1, \psi_2, \ldots\}$ ist. Die b_{pq} $(p, q = 1, 2, \ldots)$ sind so definiert:*

Wir teilen die Zahlenreihe folgendermaßen ein: Das Intervall $m\cdot 2^{k+1}+1, \ldots, (m+1)\cdot 2^{k+1}$ ist eine Zelle, $m\cdot 2^{k+1}+1, \ldots, (2m+1)\cdot 2^k$ und $(2m+1)\cdot 2^k+1, \ldots, (m+1)\cdot 2^{k+1}$ ihre erste bzw. zweite Hälfte, $(2m+1)\cdot 2^k$ ihre Mitte, k ihr Grad. Jede Zahl $p = 1, 2, \ldots$ ist dann Mitte einer einzigen Zelle, ihrer Zelle.

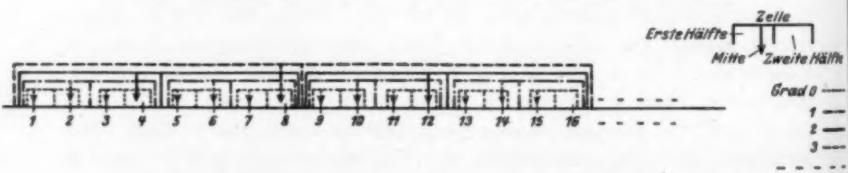
Es ist:

$$b_{pq}\begin{cases} = \quad 0 & \text{wenn von den Zellen von } p, q \text{ keine echter Teil der anderen ist.} \\ = \pm i\cdot 2^{\frac{1}{2}(3k-l)} & \text{wenn das der Fall ist, und } k, l \text{ der Grad der kleineren bzw. größeren Zelle ist.} \end{cases}$$

Das Vorzeichen $\pm$ bestimmt sich so:

	sie liegt in der ersten Hälfte	*sie liegt in der zweiten Hälfte*
p-s Zelle ist kleiner	$-$	$+$
q-s Zelle ist kleiner	$+$	$-$

(Man beachte: B ist i-mal eine schiefsymmetrische Matrix.)



3. Eine bessere Übersicht gewinnen wir über $\bar{R}$, wenn wir es in Funktionenräumen realisieren. So sei Ω das Intervall $0, 1$, wir bilden seinen Funktionenraum: die Menge aller (nach Lebesgue meßbaren) Funktionen $f(x)$ des Intervalles $0 < x < 1$, mit endlichem $\int_0^1 |f(x)|^2 dx$ — nach Anhang I ist dies ein Hilbertscher Raum, isomorph $\mathfrak{H}$. Er heiße $\mathfrak{H}^*$.

Die $e^{2n\pi i\cdot x}$, $n = 0, \pm 1, \pm 2, \ldots$ bilden bekanntlich ein vollst. norm. orth. System in $\mathfrak{H}^*$. Sei $\mathfrak{J}$ die von den $e^{2n\pi i\cdot x}$, $n = 0, 1, 2, \ldots$ aufge-

spannte abg. lin. M., $\mathfrak{J}$ ist unendlichvieldimensional, also ebenfalls ein Hilbertscher Raum (in diesem $\mathfrak{J}$ wollen wir $\bar{R}$ realisieren).

$$Uf(x) = e^{2\pi i \cdot x} f(x)$$

ist ein längentreuer (ja unitärer) Operator in $\mathfrak{H}^*$, der $\mathfrak{J}$ auf ein Teil von sich selbst abbildet. Wenn wir $\varphi_\nu(x) = e^{2\nu\pi i \cdot x}$ $(\nu = 0, 1, 2, \ldots)$ setzen so sind die $\varphi_0, \varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System in $\mathfrak{J}$, und es ist $U\varphi_\nu = \varphi_{\nu+1}$; d. h. in $\mathfrak{J}$ stimmt U mit $\bar{U}$ überein, und $\bar{R}$ ist seine Cayleysche Transformierte. Also: $\bar{R}f(x)$ hat Sinn ($f(x)$ aus $\mathfrak{J}$!) wenn $f(x) = \varphi(x) - U\varphi(x) = (1 - e^{2\pi i \cdot x})\varphi(x)$, $\varphi(x)$ aus $\mathfrak{J}$, ist, und zwar ist es dann $= i(\varphi(x) + U\varphi(x)) = i(1 + e^{2\pi i \cdot x})$. Oder auch: $\bar{R}f(x)$ hat Sinn, wenn $f(x)$, $\frac{f(x)}{1 - e^{2\pi i \cdot x}}$ beide zu $\mathfrak{J}$ gehören, und zwar ist es dann $= i\frac{1 + e^{2\pi i \cdot x}}{1 - e^{2\pi i \cdot x}} f(x) = -\operatorname{ctg.} \pi x \cdot f(x)$.

$f(x)$ gehöre zu $\mathfrak{J}$. Daß $\frac{f(x)}{1 - e^{2\pi i \cdot x}}$ zu $\mathfrak{J}$ gehört, bedeutet: sein Absolutwertquadratintegral ist endlich, und es ist zu allen $e^{2\pi i \cdot nx}$, $n = -1, -2, \ldots$ orth. Hieraus folgt die Endlichkeit von $\int\limits_0^1 \operatorname{ctg}^2 \pi x \; |f(x)|^2 dx$ (weil ja $\bar{R}f(x) = -\operatorname{ctg} \pi x \; f(x)$ ist), und wir wollen zeigen, daß diese umgekehrt zur Folge hat, daß $\frac{f(x)}{1 - e^{2\pi i \cdot x}}$ zu $\mathfrak{J}$ gehört ($f(x)$ aus $\mathfrak{J}$!).

Zunächst ist wegen $\left|\frac{1}{1 - e^{2\pi i \cdot x}}\right|^2 = \left|\frac{\operatorname{ctg} \pi x + i}{2i}\right|^2 \leqq \frac{1}{2} \operatorname{ctg}^2 \pi x + \frac{1}{2}$ auch $\int\limits_0^1 \left|\frac{f(x)}{1 - e^{2\pi i \cdot x}}\right|^2 dx$ endlich, also alle $\int\limits_0^1 \left|\frac{f(x)}{1 - \varrho e^{2\pi i \cdot x}}\right|^2 dx$ $(0 < \varrho < 1)$ absolut gleichmäßig integrierbar. Also ist für $0 < \varrho < 1$, $\varrho \to 1$

$$\int\limits_0^1 \frac{f(x)}{1 - \varrho e^{2\pi i \cdot x}} e^{-2n\pi i \cdot x} dx \to \int\limits_0^1 \frac{f(x)}{1 - e^{2\pi i \cdot x}} e^{-2n\pi i \cdot x} dx.$$

Wir haben noch zu zeigen, daß die rechte Seite für alle $n = -1, -2, \ldots$ verschwindet, es genügt also dies für die linke (und alle $0 < \varrho < 1$!) zu beweisen. Nun ist $\frac{1}{1 - \varrho e^{2\pi i \cdot x}} = \sum\limits_{\nu=0}^{\infty} \varrho^\nu e^{2\nu\pi i \cdot x}$, und die Reihe konvergiert gleichmäßig in x (ϱ fest!), daher ist

$$\sum_{\nu=0}^{N} \varrho^\nu \int\limits_0^1 f(x) e^{-2(n-\nu)\pi i \cdot x} dx = \int\limits_0^1 f(x) \Big(\sum_{\nu=0}^{N} \varrho^\nu e^{2\nu\pi i \cdot x}\Big) e^{-2n\pi i \cdot x}$$

$$\to \int\limits_0^1 \frac{f(x)}{1 - \varrho e^{2\pi i \cdot x}} e^{-2n\pi i \cdot x} dx$$

(für $N \to \infty$). Da $n = -1, 2, \ldots$, $\nu = 0, 1, 2, \ldots$ also $n - \nu = -1, -2, \ldots$

ist und $f(x)$ selbst zu $\mathfrak{J}$ gehört, ist die linke Seite stets 0, also verschwindet auch die rechte.

Nun können wir beweisen:

Satz 2**. *$\mathfrak{H}^*$ sei der soeben beschriebene Hilbertsche Funktionenraum, $\mathfrak{J}_k^{\pm}$ die von allen $e^{2n\pi i\cdot x}$, $n \geq k$ bzw. $\leq k$ $(n = 0, \pm 1, \pm 2, \ldots,$ ebenso $k)$ aufgespannte abg. lin. M. — die $\mathfrak{J}_k^{\pm}$ sind unendlichvieldimensional, also wieder Hilbertsche Räume.*

R sei der folgende Operator in $\mathfrak{H}^*$: wenn $f(x)$, $-\operatorname{ctg}\pi x\cdot f(x)$ zu $\mathfrak{H}^*$ gehören, $\left(\text{d. h. } \int_0^1 |f(x)|^2\cdot dx,\ \int_0^1 \operatorname{ctg}^2\pi x\cdot |f(x)|^2\cdot dx \text{ endlich sind}\right)$, so sei $Rf(x)$ sinnvoll, und zwar $= -\operatorname{ctg}\pi x\cdot f(x)$.

Wenn $f(x)$ in einem der $\mathfrak{J}_k^{\pm}$ liegt und $Rf(x)$ Sinn hat, so liegt auch $Rf(x)$ in diesem $\mathfrak{J}_k^{\pm}$; so können wir R in jedem $\mathfrak{J}_k^{\pm}$ zu einem H. O. dortselbst einschränken[77]*. In den $\mathfrak{J}_k^{+}$ erhalten wir so $\bar{R}$, in den $\mathfrak{J}_k^{-}$ $-\bar{R}$.*

Beweis. Für $\mathfrak{J}_0^+ = \mathfrak{J}$ haben wir dies bereits bewiesen. Die unitäre Abbildung $f(x) \Rightarrow e^{2k\pi i\cdot x}\cdot f(x)$ führt $\mathfrak{H}^*, R, \mathfrak{J}_0^+$ in $\mathfrak{H}^*, R, \mathfrak{J}_k^+$ über, die unitäre Abbildung $f(x) \Rightarrow f(1-x)$ $\mathfrak{H}^*, R, \mathfrak{J}_{-k}^+$ in $\mathfrak{H}^*, -R, \mathfrak{J}_k^-$ — also gelten die Behauptungen für alle $\mathfrak{J}_k^{\pm}$.

Man zeigt übrigens leicht, daß dieses R in $\mathfrak{H}^*$ ein hypermax. H. O. ist.

4. Wir können jedem $f(x) = \sum_{n=0}^{\infty} a_n e^{2n\pi i\cdot x}$ $\left(\sum_{n=0}^{\infty} |a_n|^2 \text{ endlich}\right)$ von $\mathfrak{J}_0^+$[78], die $\left(\text{wegen } \sum_{n=0}^{\infty} |a_n|^2 \text{ endlich, also } a_n \to 0\right)$ im Einheitskreise analytische Funktion $F(z) = \sum_{n=0}^{\infty} a_n z^n$ zuordnen. Da

$$\sum_{n=0}^{\infty} |a_n|^2 = \lim_{r\to 1} \sum_{n=0}^{\infty} r^n |a_n|^2 = \lim_{r\to 1} \int_{|z|=r} |F(z)|^2 \frac{dz}{z}$$

ist, bilden die im Einheitskreise analytischen Funktionen $F(z)$ mit beschränktem $\int_{|z|=r} |F(z)|^2 \frac{dz}{z}$ $(r \to 1)$ einen Hilbertschen Raum. Wegen $-\operatorname{ctg}\pi x = i\frac{e^{2\pi i\cdot x}+1}{e^{2\pi i\cdot x}-1}$ ist hierin $\bar{R}$ durch $\bar{R}\,F(z) = i\frac{z+1}{z-1}\,F(z)$ definiert (falls dieses wieder zur genannten Klasse gehört). — Eine andere be-

---

[77]) Aber dies ist noch keineswegs Reduzierbarkeit von R durch $\mathfrak{J}_k^{\pm}$! Denn mit Rf braucht ja nicht auch $RP_{\mathfrak{J}_k^{\pm}}f$ sinnvoll zu sein.

[78]) Die Reihe $\sum_{n=0}^{\infty} a_n e^{2n\pi i\cdot x}$ konvergiert (im Sinne der von uns stets verwendeten Metrik) „im Mittel": $\int_0^1 \left| f(x) - \sum_{n=0}^{N} a_n e^{2n\pi i\cdot x}\right|^2 dx \to 0$ (für $N \to \infty$); natürlich nicht notwendig „punktweise".

merkenswerte Realisation von $\bar{R}$ gewinnen wir, wenn wir die im Intervalle $0, \infty$ definierten Funktionen $f(x)$ $\left(\text{mit endlichem } \int_0^\infty |f(x)|^2 dx\right)$ betrachten: es entspricht dann dem Differentialoperator $i\frac{d}{dx}\ldots$ mit der Randbedingung $f(0)=0$. Wir gehen aber hierauf nicht näher ein [79]).

Es bleibt noch übrig, die bei Satz 40 benützten Eigenschaften von $\bar{R}$ zu verifizieren. Erstens: $\bar{R}$ ist nicht überall sinnvoll und nimmt nicht jeden Wert an. Wir realisieren es nach Satz 2^{**} in $\mathfrak{J}_0^+$, da $\int_0^1 \operatorname{ctg}^2 \pi x\, dx$, $\int_0^1 \operatorname{tg}^2 \pi x\, dx$ beide $=\infty$ sind, ist $\bar{R}$ für $f(x)=1$ sinnlos, und nimmt diesen Wert nie an. Zweitens: $\bar{R}$ ist weder nach oben noch nach unten halbbeschränkt. Wir realisieren es nach Satz 1^{**} und setzen $f=\psi_{2^\nu} \pm i\psi_{2^{\nu+1}}$, dann ist $(f, \bar{R}f) = \mp 2^{\nu+\frac{1}{2}}$ und $|f|^2=2$ — d. h. $(f, \bar{R}f) > C\cdot|f|^2$ bzw. $< C\cdot|f|^2$ bei beliebigem C erreicht, wenn $2^{\nu-\frac{1}{2}} > |C|$ gewählt wird.

---

[79]) Sie wird in der in Anm. [68]) genannten Arbeit erörtert werden.

(Eingegangen am 15. 12. 1928.)

Über den Cauchyschen Integralsatz für reelle und komplexe Funktionen.

Von

J. Ridder in Baarn (Niederlande).

Es ist unsere Absicht, im folgenden zwei Hauptsätze [A und B] herzuleiten, welche Bedingungen enthalten, unter denen der Cauchysche Integralsatz bei reellen Funktionen seine Gültigkeit behält bzw. unter denen sich die Analytizität einer komplexen Funktion zeigen läßt. Weiter wird gezeigt, daß die verschiedenen Funktionengruppen, welche den meistbekannten, hinreichenden Bedingungen für die Gültigkeit jener Sätze genügen, dadurch auch implizit den Bedingungen des entsprechenden Hauptsatzes genügen. Damit ist dann aufs neue gezeigt, daß jene Bedingungen hinreichend sind. Sie sind dabei möglichst weit verallgemeinert, während die in den §§ 15 und 16 enthaltenen Bedingungen neu hinzugefügt wurden.

§ 1.

In unseren Beweisen brauchen wir den Begriff der Intervallfunktion in der Ebene. Eine derartige Funktion $\Phi(J)$ ist definiert in jedem Intervall J, das einem bestimmten Gebiete der Ebene angehört und dessen Seiten den Koordinatenachsen parallel laufen. Sie heißt *beschränkt* additiv, wenn $\Phi(J) = \Phi(J_1) + \Phi(J_2) + \ldots + \Phi(J_n)$ ist, wobei das Intervall J in eine *endliche* Anzahl von Teilintervallen $J_1, J_2, \ldots, J_n$ zerfällt [1]). — Konvergiert $\Phi(J)$ in einem Punkte (x, y) des Bereiches immer nach Null, falls J eine Reihe den Punkt als inneren oder Randpunkt enthaltender und sich in den Punkt zusammenziehender Intervalle durchläuft, deren Maß nach Null konvergiert, so heißt die Intervallfunktion stetig in (x, y). — Die obere und untere Derivierte in (x, y), $\overset{+}{D}_{(x,y)}\Phi(J)$ bzw. $\overset{-}{D}_{(x,y)}\Phi(J)$, sollen nun

[1]) Wo im folgenden „additiv" steht, meinen wir „beschränkt additiv".

definiert werden als $\limsup_{m(J)=0} \frac{\Phi(J)}{m(J)}$ und $\liminf_{m(J)=0} \frac{\Phi(J)}{m(J)}$, wobei J ein *Quadrat* darstellt, das (x, y) im Innern oder auf dem Rande enthält[*]). Sind sie einander gleich, so existiert in (x, y) eine Ableitung $D_{(x,y)}\Phi(J)$.

§ 2.

Wenn in einem abgeschlossenen Intervall J, dessen Seitenlängen ein rationales Verhältnis haben, eine additive Intervallfunktion definiert ist, deren untere Derivierte in jedem Punkte > 0 ist, so ist $\Phi(J) \geqq 0$. Ist das Verhältnis der Seitenlängen irrational, so gilt dieselbe Aussage, wenn man noch Stetigkeit der Intervallfunktion in mindestens einem Punkte des abgeschlossenen J annimmt.

Im Falle eines Quadrates führte die Annahme $\Phi(J) < 0$, bei fortgesetzter Vierteilung in Quadrate, zu einem Punkte (x, y) in J, der Randpunkt oder innerer Punkt einer Folge von Quadraten Q_n wäre, deren Maß nach Null konvergierte und für die $\Phi(Q_n) < 0$ wäre. Also würde die untere Derivierte in (x, y) $\leqq 0$ sein entgegen der Voraussetzung.

Ein Intervall mit rationalem Seitenverhältnisse läßt sich in eine endliche Anzahl von Quadraten teilen. Aus der Additivität folgt damit auch in diesem Falle $\Phi(J) \geqq 0$.

Ist schließlich das Seitenverhältnis von J irrational und (x, y) ein Stetigkeitspunkt von Φ, so läßt sich J in zwei oder vier Intervalle I_j teilen [ausgenommen wenn (x, y) Eckpunkt von J ist] mittels der beiden Achsenparallelen durch (x, y). Jedes I_j [oder J selbst] ist wieder zu teilen in eine endliche Anzahl von Quadraten nebst einem Intervall U_j [oder U], das *kein* Quadrat zu sein braucht und (x, y) zum Eckpunkte hat. Dabei kann zufolge der Stetigkeit von Φ in (x, y) jedes Intervall U_j [oder U] so gewählt werden, daß bei positivem ε:

$$|\Phi(U_j)| < \varepsilon \quad [\text{oder } |\Phi(U)| < \varepsilon]$$

ist. Somit ist in jedem Falle:

$$\Phi(J) > -4\varepsilon$$

und, da ε beliebig klein sein darf, ist auch

$$\Phi(J) \geqq 0.$$

Die Voraussetzung $D^-\Phi > 0$ darf durch $D^-\Phi \geqq 0$ ersetzt werden.

---

*) Ist Φ definiert im abgeschlossenen Bereiche B, welcher aus einem Gebiete G hervorgeht durch Hinzufügung seiner Randpunkte, so lassen sich die oberen und unteren Derivierten in einem Randpunkte (x, y) definieren mittels der in B liegenden und (x, y) als Randpunkt enthaltenden Quadrate. Auch die Definition der Stetigkeit in einem Randpunkte läßt sich auf analoge Weise geben.

Durch Anwendung des vorherigen Satzes auf die Funktion $\Phi(J)+\varepsilon\cdot m(J)$, wobei ε positiv und beliebig klein, folgt:

$$\Phi(J)+\varepsilon\cdot m(J)\geqq 0.$$

Also für $\lim \varepsilon = 0$ wird:

$$\Phi(J)\geqq 0.$$

§ 3.

Im folgenden wird eine positive, stetige, additive Intervallfunktion konstruiert, deren Derivierten in einer Menge vom Maße Null positiv unendlich werden.

Sei M die ganz beliebige Menge vom Maße Null. Sie läßt sich einschließen in eine Folge von offenen Mengen $O_1 > O_2 > \ldots$, so daß $m(O_k) < \varepsilon_k$ und $\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_k+\ldots<\varepsilon$ ist. Die charakteristische Funktion von O_k sei h_k. Dann ist im Intervall J:

$$\iint\limits_J h_k\,dx\,dy = m(O_k\cdot J)\leqq m(O_k) < \varepsilon_k.$$

Für jeden Punkt der Menge O_k ist dann die Ableitung dieses Integrals $=1$. Setzen wir $\chi_k(J)=\sum\limits_{\nu=1}^{k}\iint\limits_J h_\nu\,dx\,dy$, so ist in O_k: $D_{(x,y)}\chi_k(J)=k$. Da $\chi_k(J)<\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_k<\varepsilon$ ist und also $\lim\limits_{k=\infty}\chi_k(J)$ endlich ist, ist nach einem Satze von Levi [3]):

$$\lim_{k=\infty}\chi_k(J)=\lim_{k=\infty}\iint\limits_J\sum_{\nu=1}^{k}h_\nu\,dx\,dy=\iint\limits_J\lim_{k=\infty}\sum_{\nu=1}^{k}h_\nu\,dx\,dy<\varepsilon$$

für jedes Intervall J. Setzen wir $\chi(J)=\lim\limits_{k=\infty}\chi_k(J)$ und $H=\lim\limits_{k=\infty}\sum\limits_{1}^{k}h_\nu$, so ist

$$\chi(J)=\iint\limits_J H\,dx\,dy<\varepsilon$$

die gesuchte Funktion. Denn in M ist $D^{\pm}_{(x,y)}\chi(J)>D_{(x,y)}\chi_k(J)=k$ für jedes ganze, positive k. Weiter ist $\chi(J)$ sogar totalstetig und absolut additiv.

§ 4.

Wenn in einem abgeschlossenen Intervall J, dessen Seitenlängen ein rationales Verhältnis haben, eine additive Intervallfunktion definiert ist, deren untere Derivierte in keinem Punkte $-\infty$ und fast überall $\geqq 0$ ist, so ist auch $\Phi(J)\geqq 0$. Ist das Verhältnis der Seitenlängen irrational, so gilt dieselbe Aussage, wenn man noch Stetigkeit der Intervallfunktion in einem Punkte des abgeschlossenen J annimmt.

---

[3]) Siehe z. B. Schlesinger und Pleßner, Lebesguesche Integrale, Nr. 27 oder Carathéodory, Vorles. über reelle Funktionen, § 383, Satz 4.

M sei die Teilmenge in J, von der man nicht weiß, ob in ihren Punkten $D^- \Phi \geqq 0$ ist. Nun sei $\chi(J)$ das oben angedeutete Beispiel einer positiven, stetigen und additiven Intervallfunktion, deren Derivierten in der Menge M unendlich werden und deren Wert in jedem Intervall $< \varepsilon$ ist. Betrachten wir $\Phi_1(J) = \Phi(J) + \chi(J)$, so ist in allen Punkten des abgeschlossenen Intervalls $D^- \Phi_1 \geqq 0$. Mit § 2 folgt $\Phi_1(J) \geqq 0$, also auch, da $\chi(J)$ willkürlich klein ist, $\Phi(J) \geqq 0$.

§ 5.

Wir ändern die vorangehende Definition der Stetigkeit in der Weise ab, daß wir die Intervallfunktion Φ in jedem Punkte (x, y) eines Gebietes G als stetig betrachten, wenn immer $\lim\limits_{m(J)=0} \Phi(J) = 0$ ist, falls J eine Reihe den Punkt als inneren oder Randpunkt enthaltender Intervalle durchläuft, deren Maß nach Null konvergiert. Das Intervall braucht sich also nicht immer in den Punkt (x, y) zusammenzuziehen[4]; Die vorhergehenden Sätze gelten auch bei dieser Definition ungeändert. Wir können nun hinzufügen:

Ist Φ eine im abgeschlossenen Intervall J definierte, additive Intervallfunktion, welche in J (im abgeänderten Sinne) stetig ist, so ist $\Phi(J) \geqq 0$, falls in J:* 1. *die untere Derivierte nur in einer abzählbaren Menge E $-\infty$ werden kann,* 2. *die untere Derivierte fast überall $\geqq 0$ ist.*

Es sei E_n die Vereinigungsmenge aller derjenigen Punkte des abgeschlossenen Intervalls J, welche in J liegenden, abgeschlossenen Quadraten J_n angehören, für die $\frac{\Phi(J_n)}{m(J_n)} < -n$ ist. Hierbei sei n eine positive ganze Zahl. Aus der Stetigkeit* von Φ folgt, daß jeder Randpunkt eines J_n, welcher zugleich innerer Punkt von J ist, auch innerer Punkt eines Quadrates J_n' ist, das 1. J_n enthält, 2. im abgeschlossenen J liegt, und 3. für das ebenfalls $\frac{\Phi(J_n')}{m(J_n')} < -n$ ist. Jeder Punkt (x, y) von E_n, welcher zugleich Randpunkt von J ist, ist immer Randpunkt eines Quadrates J_n in J, das so gewählt werden kann, daß (x, y) kein Eckpunkt von J_n ist. Dieser Punkt ist dadurch auch als innerer Punkt eines Intervalls I_n zu betrachten, welches aus diesem J_n hervorgeht durch Verschiebung der durch (x, y) laufenden Seite um einen Abstand $\frac{1}{n}$ nach der Außenseite des Intervalles J. Die aus den inneren Punkten der J_n, (J_n') und I_n existierenden Menge E_n' ist offen. Der Durchschnitt aller

[4]) Die neue Stetigkeitsdefinition in einem inneren Punkte oder Randpunkte eines abgeschlossenen Bereiches läßt sich auf analoge Weise geben.

dieser offenen Mengen E_n' fällt zusammen mit dem Durchschnitte D aller Mengen E_n. Somit ist D eine innere Grenzmenge.

Nun ist die abzählbare Menge E derjenigen Punkte des abgeschlossenen J, wo $D^- \Phi = -\infty$ ist, mit D identisch. Denn zu jedem Punkte (x, y) des Durchschnittes D gehört ein Intervall J_n aus J, das (x, y) im Innern oder auf dem Rande enthält und so daß $\frac{\Phi(J_n)}{m(J_n)} < -n$ oder $|\Phi(J_n| > n \cdot m(J_n)$ ist. Wäre $\liminf_{n=\infty} m(J_n)$ eine positive Zahl δ, so wäre $\lim_{n=\infty} |\Phi(J_n)| = +\infty$. Es existierte dann eine Folge von Quadraten J_ν aus den J_n und ein in J liegendes Intervall I derartig, daß 1. die Seiten der J_ν mit zunehmendem ν konvergierten gegen die übereinstimmenden Seiten von I, 2. $m(I) \geq \delta$ und 3. zufolge der Stetigkeit * $|\Phi(I)| = \lim_{\nu=\infty} |\Phi(J_\nu)| = +\infty$ wäre. Da nur endliche Werte von Φ zugelassen sind, ist dies unmöglich. Also ist $\liminf_{n=\infty} m(J_n) = 0$ und gehört dadurch (x, y) zu E. Daß umgekehrt jeder Punkt von E zum Durchschnitt der E_n gehört, ist evident.

Da E abzählbar und innere Grenzmenge ist, ist sie nirgends dicht auf jeder perfekten Menge in J.[5])

Die Menge A der Punkte von J, in deren jeder Umgebung Intervalle I liegen mit $\Phi(I) < 0$, ist abgeschlossen. Es sei (x, y) ein Punkt von A und I ein Intervall aus einer Umgebung von (x, y), so daß $\Phi(I) < 0$ ist. Die Achsenparallelen durch (x, y) teilen I in zwei oder vier Intervalle I_j [ausgenommen, wenn (x, y) Eckpunkt oder äußerer Punkt von I ist]. Aus der Additivität von Φ folgt, daß für eines der I_j: $\Phi(I_j) < 0$ ist. Aus der Stetigkeit* von Φ folgt weiter, daß I_j [oder I] sich durch hinreichend kleine Verschiebung überführen läßt in ein Intervall I_j' [oder I'], das (x, y) nicht enthält und so daß auch $\Phi(I_j') < 0$ [oder $\Phi(I') < 0$] ist. Fortgesetzte Vierteilung von I_j' [oder I'] zeigt schließlich, daß dieses Intervall einen Punkt von A enthält. Da (x, y) Intervalle I mit $\Phi(I) < 0$ und dadurch auch Punkte von A in einer jeden Umgebung hat, ist (x, y) Häufungspunkt von A. Somit ist A sogar perfekt.

Aus der Definition von A folgt, daß in der Komplementärmenge $(J - A)$: $D^- \Phi \geq 0$ ist. E muß dadurch eine Untermenge von A sein.

E ist auf A nirgends dicht. A enthält innere Punkte von J. Es existiert dadurch ein offenes Unterintervall U von J, so daß die Menge $(U \cdot A)$ nicht leer ist und auf $(U \cdot A)$: $D^- \Phi \neq -\infty$ ist. In jeder beliebig kleinen Umgebung V eines Punktes von $(U \cdot A)$ ist dann $D^- \Phi \neq -\infty$ auf $(V \cdot A)$ und ≥ 0, also auch $\neq -\infty$, auf der Komplementärmenge

---

[5]) Siehe z. B. Carathéodory, Reelle Funktionen, § 81.

$(V - V \cdot A)$. Nach § 4 wäre somit für jedes Intervall dieser Umgebung $\Phi \geqq 0$ entgegen unserer Annahme. Die Menge A ist leer.

Anwendung auf Φ und $-\Phi$ liefert:

Die im abgeschlossenen Intervall J definierte, stetige, additive Intervallfunktion Φ ist in J identisch Null, falls* 1. *die unteren und oberen Derivierten nur in einer abzählbaren Menge unendlich werden können,* 2. *fast überall beide Derivierten Null sind*[6]).

§ 6.

Das Linienintegral: $\int_J p(x,y)\,dx + q(x,y)\,dy$ ist ein Beispiel einer additiven Intervallfunktion. Anwendung des letzten Satzes liefert:

Hauptsatz A: *Es seien $p(x,y)$ und $q(x,y)$ in einem bestimmten (beschränkten) Gebiete G der Ebene nach x bzw. y linear summierbar. Dann ist $\Phi(J) = \int_J p\,dx + q\,dy$ in G identisch Null, falls in G* 1. *$\Phi(J)$ stetig* ist,* 2. *$D^+\Phi$ und $D^-\Phi$ nur in einer abzählbaren Menge unendlich werden können,* 3. *fast überall $D\Phi = 0$ ist*[7]).

Sind p und q (zweidimensional) stetig nach x und y, so braucht man nur 2. *und* 3. *anzunehmen.* Denn in jedem in G liegenden, abgeschlossenen Intervall J sind dann p und q beschränkt und gleichmäßig stetig; daraus folgert man leicht, daß Φ stetig* ist in J und dadurch $\equiv 0$ in J, — also auch in G. *Außerdem verschwindet in diesem Falle das Linienintegral über jede einfache, geschlossene, rektifizierbare Kurve in G.*[8])

§ 7.

Verallgemeinerung eines Satzes von Lichtenstein: *In einem (beschränkten) Gebiete G der Ebene ist $\Phi(J) = \int_J p\,dx + q\,dy$ identisch Null, falls p und q diesen Bedingungen genügen:*

---

[6]) *Statt* 2. genügt auch: *$D^+\Phi$ [oder $D^-\Phi$] ist fast überall Null.* — Für den Beweis und für weitere Sätze über additive Intervallfunktionen in der Ebene siehe einen im Nieuw Archief voor Wiskunde (Amsterdam) (2) **16**, 1 (1929) erschienenen Artikel.

[7]) *Statt* 3. genügt auch: *$D^+\Phi$ [oder $D^-\Phi$] ist fast überall Null in G.* — Vgl loc. cit. [6]), § 8.

[8]) Siehe für den Beweis des letzten Teiles: Pollard, Proc. of the London Math. Soc. (2) **21** (1923), § 8 (B) and (C). Dort wird bewiesen: „Wenn in einem beschränkten Gebiete G das Linienintegral einer stetigen Funktion $f(z)$, wobei $z = x + iy$, über den Rand eines jeden abgeschlossenen Intervalls verschwindet, so gilt dasselbe über jede einfache, geschlossene, rektifizierbare Kurve in G." Die Übertragung des Beweises für den im Texte gegebenen, reellen Fall bietet keine Schwierigkeiten.

1. *sie sind in jedem in G liegenden, abgeschlossenen Bereiche beschränkt und in G in bezug auf jede einzelne Variable stetig,*

2. *„mit Ausnahme einer höchstens abzählbaren Menge E" existiert für jeden Punkt von G eine bestimmte Umgebung folgender Art:*

$$\frac{1}{h}[\delta_x q(x, y) - \delta_y p(x, y)]$$

ist für alle Wertepaare (x, y) aus dieser Umgebung „jedoch bei Vernachlässigung einer Menge vom Maße Null" und alle Werte von $h > 0$ dem absoluten Betrage nach unterhalb einer endlichen Schranke,

3. *fast überall in G ist*

$$\lim_{h=0} \frac{1}{h}[\delta_x q(x, y) - \delta_y p(x, y)] = 0, \qquad h > 0.$$

Hierbei ist

$$\delta_x q(x, y) = q(x+h, y) - q(x, y) \text{ und } \delta_y p(x, y) = p(x, y+h) - p(x, y).$$

In jedem in G liegenden, abgeschlossenen Intervall I ist nach 1. $p(x, y)$ beschränkt. Es sei A die obere Schranke und (ξ, η) ein willkürlicher Punkt in I. Dann ist für jedes abgeschlossene Intervall J, das in I liegt und dessen untere oder obere Seite (mit der Länge h) durch (ξ, η) geht:

$$(1) \qquad \left|\int_J p\,dx\right| \leqq A \times 2h.$$

Also konvergiert $\int_J p\,dy$ nach Null, wenn h willkürlich klein wird.

Für jedes abgeschlossene Intervall J_n $[x_1 \leqq x \leqq x_2;\ \eta \leqq y \leqq d_n$ oder $d_n \leqq y \leqq \eta]$ in I, dessen Horizontalseite mit Ordinate d_n sich mit zunehmendem n der anderen Seite nähert, ohne daß diese in Länge oder Lage sich ändert, folgt durch Anwendung eines Satzes von Lebesgue (über Grenzübergang unter dem Integralzeichen bei gleichmäßig beschränkten Funktionen):

$$\lim_{d_n=\eta} \int_{x_1}^{x_2} p(x, d_n)\,dx = \int_{x_1}^{x_2} \lim_{d_n=\eta} p(x, d_n)\,dx = \int_{x_1}^{x_2} p(x, d)\,dx.$$

Also:

$$(2) \qquad \lim_{d_n=\eta} \int_{J_n} p\,dx = 0.$$

Aus dem Beweise des Lebesgueschen Satzes[*]) geht weiter hervor, daß das in (2) enthaltene Integral sich mit zunehmendem n der Null gleichmäßig nähert, wenn man als Integrationsgrenzen alle möglichen x_1 und x_2 zuläßt, welche der Bedingung: $a \leqq x_1 < x_2 \leqq b$ genügen.

---

[*]) Siehe z. B. de la Vallée Poussin, Intégrales de Lebesgue, etc. (Coll. Borel), § 39.

Hierdurch und mit (1) folgert man leicht, daß $\int_J p\,dx$ stetig* ist in I. Dasselbe gilt von $\int_J q\,dy$; also auch von $\Phi(J)$.

Aus 1. folgt, daß $p(x,y)$ und $q(x,y)$ zur ersten Baireschen Klasse gehören[10]) und somit in jedem abgeschlossenen Bereiche in G summierbar sind. Nach dem Satze von Lebesgue folgt aus 2. und 3. für jedes Intervall J in einer bestimmten Umgebung eines jeden Punktes (ξ,η) „mit Ausnahme der höchstens abzählbaren Menge E":

$$
\begin{aligned}
&\lim_{h=0}\frac{1}{h}\iint_J[\delta_x q(x,y)-\delta_y p(x,y)]\,dx\,dy \\
(3)\qquad &=\iint_J\lim_{h=0}\frac{1}{h}[\delta_x q(x,y)-\delta_y p(x,y)]\,dx\,dy=0.
\end{aligned}
$$

Betrachten wir das Integral:

$$
\begin{aligned}
\iint_J\frac{1}{h}[\delta_x q(x,y)-\delta_y p(x,y)]\,dx\,dy&=\frac{1}{h}\iint_J[q(x+h,y)-q(x,y)]\,dx\,dy\\
-\frac{1}{h}\iint_J[p(x,y+h)-p(x,y)]\,dx\,dy&=\frac{1}{h}\iint_{J_1}q(x,y)\,dx\,dy-\frac{1}{h}\iint_{J_2}q(x,y)\,dx\,dy\\
-\frac{1}{h}\iint_{J_3}p(x,y)\,dx\,dy&+\frac{1}{h}\iint_{J_4}p(x,y)\,dx\,dy.
\end{aligned}
$$

Genügen die Koordinaten von J den Ungleichungen: $\alpha\leqq x\leqq\beta$, $\gamma\leqq y\leqq\delta$, so gilt für J_1: $\beta\leqq x\leqq\beta+h$, $\gamma\leqq y\leqq\delta$; für J_2: $\alpha\leqq x\leqq\alpha+h$, $\gamma\leqq y\leqq\delta$; für J_3: $\alpha\leqq x\leqq\beta$, $\delta\leqq y\leqq\delta+h$; und für J_4: $\alpha\leqq x\leqq\beta$, $\gamma\leqq y\leqq\gamma+h$.

Nach einem Satze von Fubini ist:

$$\frac{1}{h}\iint_{J_1}q(x,y)\,dx\,dy=\frac{1}{h}\int_\beta^{\beta+h}dx\int_\gamma^\delta q(x,y)\,dy.$$

Da $q(x,y)$ stetig nach x ist und beschränkt im abgeschlossenen Intervall J_1, ist auch $\int_\gamma^\delta q(x,y)\,dy$ stetig nach x bei $\beta\leqq x\leqq\beta+h$. Daraus folgt:

$$\lim_{h=0}\frac{1}{h}\iint_{J_1}q(x,y)\,dx\,dy=\lim_{h=0}\frac{1}{h}\int_\beta^{\beta+h}dx\int_\gamma^\delta q(x,y)\,dy=\int_\gamma^\delta q(\beta,y)\,dy.$$

Wendet man denselben Grenzübergang an auf die über J_2, J_3 und J_4 er-

[10]) Siehe z. B. Carathéodory, Reelle Funktionen, § 558 Satz 2.

streckten Doppelintegrale, so liefert Addition schließlich:

$$(4) \qquad \lim_{h=0} \frac{1}{h} \iint_J [\delta_x q(x,y) - \delta_y p(x,y)]\, dx\, dy = \int_J p\, dx + q\, dy.$$

Aus (3) und (4) folgt, daß $D\,\Phi = 0$ ist in jedem Punkte von G, die Menge E vielleicht ausgenommen. Demnach liefert Anwendung des Hauptsatzes A, daß $\Phi(J) \equiv 0$ ist in G[11]).

Korollar, Verallgemeinerung eines Montelschen Satzes:

$$\Phi(J) = \int_J p\, dx + q\, dy \ \textit{ist} \equiv 0$$

im Gebiete G unter den Bedingungen:

1. *p und q sind in jedem in G liegenden, abgeschlossenen Bereiche beschränkt und in bezug auf jede einzelne Variable stetig*;

2. *„mit Ausnahme einer höchstens abzählbaren Menge E" existiert für jeden Punkt P von G eine bestimmte Umgebung folgender Art: Die oberen und unteren Derivierten von p nach y und von q nach x sind in dieser Umgebung beschränkt „bei Vernachlässigung einer abzählbaren Menge von Punkten in dieser Umgebung"* [12]);

3. *fast überall in G ist* $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$.

Nach einem Satze von Lebesgue [13]) folgt aus Beschränktheit im Sinne von 2. auch Beschränktheit jener Derivierten, wenn in der betrachteten Umgebung des Punktes P keine abzählbare Menge vernachlässigt wird. Da nach Dini bei Funktionen einer Veränderlichen die oberen und unteren Derivierten und die zugehörigen Differenzenquotienten in einem abgeschlossenen Intervall dieselben oberen und unteren Schranken haben [14]), so sind hier in der Umgebung von P die Differenzenquotienten von p nach y gleichmäßig beschränkt auf allen Parallelen zur y-Achse, also beschränkt in der ganzen Umgebung von P. Das letztere gilt auch von den Differenzenquotienten von q nach x. Damit sind alle Bedingungen des verallgemeinerten Lichtensteinschen Satzes erfüllt.

§ 8.

Satz: *Im (beschränkten) Gebiete G existieren zwei abzählbare Mengen E_1 und E_2, so daß für die in G definierten Funktionen $p(x,y)$ und $q(x,y)$ folgendes gilt*:

---

[11]) Vgl. Lichtenstein, Sitzungsber. Berliner Math. Ges. (1910), Satz III und Beweis.

[12]) Hieraus folgt Stetigkeit nach x und y von beiden Funktionen in P.

[13]) Siehe Lebesgue, Leçons sur l'intégration (Coll. Borel), p. 80.

[14]) Siehe z. B. Carathéodory, Reelle Funktionen, § 476, Satz 5.

1. *p und q sind in G linear stetig nach x und nach y;*

2. *in jedem Punkte von* $(G - E_1)$ *sind die partiellen oberen und unteren Derivierten von p nach y und von q nach x endlich;*

3. *für jeden Punkt von* $(G - E_2)$ *existiert eine Umgebung U, in der* $\frac{\partial p}{\partial y}$ *und* $\frac{\partial q}{\partial x}$ *summierbar sind auf der Teilmenge von U, auf der beide existieren;*

4. *fast überall ist* $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$ *auf derjenigen Menge in G, auf der beide existieren;*

5. $\Phi(J) = \int_J p\,dx + q\,dy$ *ist stetig* in den Punkten von* E_2.[15])

Dann ist in G: $\Phi \equiv 0$.

Da p und q stetig sind nach x und nach y, sind sie als Funktionen von zwei Veränderlichen meßbar[16]). Die Mengen K_1 und K_2 in G, auf der $\frac{\partial p}{\partial y}$ bzw. $\frac{\partial q}{\partial x}$ existiert, sind dadurch auch meßbar[16]). Da die oberen und und unteren Derivierten von p nach y auf einer Parallelen zur y-Achse nur in einer abzählbaren Menge unendlich werden können, so sind sie, nach einem Satze von Montel und Denjoy[17]), auf der Parallelen einander gleich außer in einer linearen Nullmenge. Daraus folgt, daß K_1 und G gleiches Flächenmaß haben. Dasselbe gilt von K_2 und G. Somit existieren beide Ableitungen fast überall in G gleichzeitig.

Nehmen wir $\frac{\partial p}{\partial y}$ und $\frac{\partial q}{\partial x}$ gleich Null dort wo die partiellen Ableitungen nicht existieren. Dann ist in einer bestimmten Umgebung eines jeden Punktes von $(G - E_2)$ für ein willkürliches Intervall J $[x_1 \leq x \leq x_2;\ y_1 \leq y \leq y_2]$:

$$\iint_J \frac{\partial p}{\partial y}\,dx\,dy = \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} \frac{\partial p}{\partial y}\,dy.$$

Für jeden Wert von x, für den das innere Integral existiert, ist, da $\frac{\partial p}{\partial y}$ höchstens abzählbar viele Unendlichkeitsstellen hat und p stetig ist nach y:

$$\int_{y_1}^{y_2} \frac{\partial p}{\partial y}\,dy = p(x, y_2) - p(x, y_1)\ {}^{18)}.$$

---

[15]) Diese Bedingung darf man fortlassen, wenn unter 1. (zwei-dimensionale) Stetigkeit von p und q nach x und y angenommen wird (vgl. § 6). Dann ist außerdem das Linienintegral Null über jede einfache geschlossene rektifizierbare Kurve in G. — In der Formulierung des Satzes sind dann auch die Mengen E_1 und E_2 zusammenzufassen.

[16]) Siehe z. B. Carathéodory, Reelle Funktionen, § 558, Satz 2 und § 557, Satz 1.

[17]) Siehe z. B. Hobson, Theory of functions I, § 298.

[18]) Nach Carathéodory, Reelle Funktionen, § 527, Satz 4.

Somit ist:

$$\iint\limits_J \frac{\partial p}{\partial y}\,dx\,dy = \int\limits_{x_1}^{x_2} [p(x, y_2) - p(x, y_1)]\,dx,$$

oder da $\int\limits_{x_1}^{x_2} p(x, y_2)\,dx$ und $\int\limits_{x_1}^{x_2} p(x, y_1)\,dx$ nach 1. existieren, auch:

$$\iint\limits_J \frac{\partial p}{\partial y}\,dx\,dy = -\int\limits_J p\,dx. \tag{5}$$

Auf dieselbe Weise läßt sich zeigen:

$$\iint\limits_J \frac{\partial q}{\partial x}\,dx\,dy = \int\limits_J q\,dy. \tag{6}$$

Nach 4. liefert Subtraktion von (5) und (6):

$$\int\limits_J p\,dx + q\,dy = 0.$$

Auf $(G - E_2)$ ist dadurch $D\Phi = 0$.

In jedem in G liegenden abgeschlossenen Intervall $J\,[x_1 \leqq x \leqq x_2;\ y_1 \leqq y \leqq y_2]$ ist Φ stetig* in den Punkten von $(J \cdot E_2)$.

Es sei (ξ, η) ein innerer Punkt von $(J - J \cdot E_2)$. Enthält die Parallele L_η zur x-Achse durch (ξ, η) keinen Punkt von $(J \cdot E_2)$, so existiert für jeden Punkt von $(L_\eta \cdot J)$ eine Umgebung, in der $\Phi \equiv 0$ ist. Daraus folgt durch Anwendung des bekannten Borelschen Lemmas, daß es ein abgeschlossenes Intervall $J_1\ [x_1 \leqq x \leqq x_2;\ y_3 \leqq y \leqq y_4]$ gibt mit $y_1 \leqq y_3 < \eta \leqq y_4 \leqq y_2$, so daß auch in J_1: $\Phi \equiv 0$ ist. Somit ist $\lim\limits_{m(I_n)=0} \Phi(I_n) = 0$, wenn I_n eine abzählbare Folge von Intervallen durchläuft, die (ξ, η) im Innern oder auf dem Rande enthalten, ganz in J liegen und deren Vertikalseiten nach Null konvergierende Längen haben.

Enthält L_η dagegen Punkte von $(J \cdot E_2)$, so ist die Durchschnittsmenge D von L_η und $(J \cdot E_2)$ abgeschlossen, da auch E_2 zufolge ihrer Definition abgeschlossen ist. Es sei (ξ_2, η) die obere Schranke derjenigen Punkte von D, welche auf der linken Seite von (ξ, η) liegen, und (ξ_3, η) die untere Schranke der Punkte von D, auf der rechten Seite von (ξ, η). Nehmen wir vorläufig die Existenz beider Teilmengen an. Dann zerfällt $(L_\eta \cdot J)$ durch (ξ_1, η) und (ξ_2, η) in drei Teile, von links nach rechts: $L_1\ [x_1 \leqq x \leqq \xi_1]$, $L_2\ [\xi_1 \leqq x \leqq \xi_2]$ und $L_3\ [\xi_2 \leqq x \leqq x_2]$.

Für eine Folge von in J liegenden Intervallen I_n, welche (ξ, η) im Innern oder auf dem Rande enthalten, deren Vertikalseiten nach Null konvergierende Längen haben und deren linke Seiten L_1 überschneiden,

ist $\lim_{n=\infty} \Phi(I_n) = 0$. Denn alle I_n enthalten als inneren Punkt oder Randpunkt (ξ_1, η). Eine auf analoge Weise definierte Folge von Intervallen U_n, welche (ξ, η) enthalten und L_η überschneiden, hat dieselbe Eigenschaft, da alle U_n (ξ_2, η) enthalten.

Eine Folge von Intervallen V_n, welche (ξ, η) enthalten, deren Vertikalseiten sich der Null nähern und deren x-Koordinaten alle im abgeschlossenen Intervall (ξ_1, ξ_2) liegen, hat auch die Eigenschaft: $\lim_{n=\infty} \Phi(V_n) = 0$. Denn bei willkürlich positivem ε existieren Umgebungen (Intervalle) W_1 und W_2 von (ξ_1, η) bzw. (ξ_2, η), in denen $|\Phi| < \varepsilon$ bleibt für diejenigen Intervalle, welche (ξ_1, η) oder (ξ_2, η) enthalten. Die abgeschlossene Strecke S_1, welche W_1 und W_2 aus L_η hinausschneiden, liegt ganz im Innern eines Intervalls W_3, in dem $\Phi \equiv 0$ ist. Daraus folgt, daß im Innern von $W_1 + W_2 + W_3$: $|\Phi| < 4\varepsilon$ ist für alle Intervalle, welche (ξ, η) enthalten. Da jedoch V_n für genügend großes n ganz im Innern von $W_1 + W_2 + W_3$ liegt und ε willkürlich positiv ist, ist die Konvergenz bewiesen.

Man folgert nun leicht, daß auch, wenn die Überschneidungen mit L_η willkürlich liegen, die Konvergenz nach Null stattfindet; so auch in den Fällen, wo D auf L_η nicht an „beiden" Seiten von (ξ, η) Punkte besitzt.

Übereinstimmende Betrachtungen gelten, wenn man Intervallfolgen um (ξ, η) betrachtet, deren Horizontalseiten nach Null konvergierende Längen besitzen.

Somit wird auch eine jede Folge von abgeschlossenen Intervallen, welche (ξ, η) enthalten, in J liegen und deren Maß nach Null konvergiert, nach Null konvergierende Werte von Φ geben. Das heißt: in J ist Φ auch stetig* in den nicht zu E_2 gehörigen inneren Punkten.

Im offenen Intervall J sind die Bedingungen des Hauptsatzes A erfüllt. Daraus folgt: $\Phi \equiv 0$ in J, also auch in G.

§ 9.

Es sei $f(z) = u(x, y) + i v(x, y)$. Dann läßt sich $\int_J f(z)\, dz$ schreiben als $\int_J u\,dx - v\,dy + i \int_J v\,dx + u\,dy = \Phi_1(J) + i\Phi_2(J)$. Wir definieren $D\Phi = D\Phi_1 + i D\Phi_2$ und betrachten Φ als stetig*, wenn Φ_1 und Φ_2 es sind. Anwendung vom Hauptsatze A auf die reellen und imaginären Teile des Linienintegrals liefert:

„Es seien $u(x, y)$ und $v(x, y)$ in einem (beschränkten) Gebiete G der Ebene nach x und nach y linear summierbar. Dann ist

$$\Phi(J) = \int_J f(z)\, dz = \Phi_1(J) + i\Phi_2(J)$$

in G identisch Null, falls in G: 1. $\Phi(J)$ stetig* ist, 2. die oberen und unteren Derivierten von Φ_1 und Φ_2 nur in einer abzählbaren Menge unendlich werden können[19]), 3. fast überall $D\Phi = 0$ ist"[20]).

Von Rademacher wurde bewiesen:

„Ist $g(z)$ eine komplexe Funktion im beschränkten Gebiete G der Ebene der $z = x + iy$, deren Real- und Imaginärteil für sich in G summierbar und auf jeder in G liegenden achsenparallelen Strecke auch linear summierbar sind, und ist für jeden achsenparallelen Rechtecksrand C, der samt seinem Innern in G liegt,

$$\int_C g(z)\,dz = 0,$$

so existiert eine analytische Funktion, welche in G fast überall mit $g(z)$ zusammenfällt.

$g(z)$ ist mit dieser analytischen Funktion identisch in G, wenn:

a) $g(z)$ stetig ist nach x oder nach y in G, oder:

b) $g(z)$ auf allen in G liegenden, abgeschlossenen Strecken, welche einer fest gewählten Achse parallel laufen, beschränkt ist und approximativ stetig als Funktion der zugehörigen Veränderlichen"[21]).

---

[19]) Oder in jedem Punkte von G, vielleicht mit Ausnahme einer abzählbaren Menge, $\limsup\limits_{m(J)=0} \frac{|\Phi(J)|}{m(J)}$ endlich ist, wobei der betrachtete Punkt im Innern oder auf dem Rande des *Quadrates* J liegt (vgl. § 1).

[20]) Statt 3. genügt: „aus einer jeden der beiden Gruppen $[D^-\Phi_1, D^+\Phi_1]$ und $[D^-\Phi_2, D^+\Phi_2]$ ist eine der Derivierten fast überall in G gleich Null". Und auch: „$\liminf\limits_{m(J)=0} \frac{|\Phi(J)|}{m(J)} = 0$ fast überall in G, wobei das *Quadrat* J sich auf (x, y) zusammenzieht (vgl. § 1)." — Vgl. für den Beweis der letzten Behauptung loc. cit. [9]) § 8 und § 19.

[21]) Siehe Rademacher, Math. Zeitschr. 4 (1919), Satz I. — Die Formulierung des Satzes ist in der zitierten Arbeit nicht ganz richtig; dort wird, ohne Bedingung a) oder b) anzunehmen, aus den „übrigen", oben im Texte gegebenen Bedingungen gefolgert: „$g(z)$ ist analytisch regulär in G". Jedoch genügt eine Funktion $g(z)$, gleich Null in allen Punkten (x, y), deren beide Koordinaten rational sind, und gleich Eins in den übrigen Punkten der Ebene, den „übrigen" Bedingungen, ohne selbst analytisch regulär zu sein; ein sehr einfaches Beispiel dieser Art ist eine Funktion, gleich Null in einem bestimmten Punkte P und gleich Eins in allen übrigen Punkten. — Rademacher zeigt, daß im Innern eines in G liegenden, abgeschlossenen Intervalls J $[x_1 \leq x \leq x_2;\ y_1 \leq y \leq y_2]$ die Funktion:

$$G(z) = G(x+iy) = \int_{x_1}^{x} g(\xi + iy_1)\,d\xi + i\int_{y_1}^{y} g(x+i\eta)\,d\eta$$

analytisch regulär ist. Die Ableitung von $G(z)$ nach z ist eine analytische Funktion, welche im allgemeinen „fast überall" in J mit $g(z)$ übereinstimmt. In den unter

(Fortsetzung der Fußnote [21]) auf nächster Seite.)

Definitionen. „Eine reelle Funktion $f(x)$, die in der Umgebung eines Punktes ξ definiert ist, heißt an der Stelle ξ approximativ stetig, wenn für jedes $\varepsilon > 0$ die Menge $\mathfrak{M}(\xi, \varepsilon)$ der Punkte x, für die

$$|f(x) - f(\xi)| < \varepsilon$$

ist, im Punkte ξ die *innere* Dichte 1 hat.

Das soll bedeuten:

$$\liminf_{m(J)=0} \frac{m_i[\mathfrak{M}(\xi, \varepsilon)\cdot J]}{m(J)} = 1,$$

wobei J ein Intervall mit Mittelpunkte ξ-, $(\mathfrak{M}\cdot J)$ den Durchschnitt der Mengen $\mathfrak{M}$ und J- und $m_i[\mathfrak{M}\cdot J]$ das innere Maß des Durchschnittes bedeutet.“ — „$g(z) = u(x, y) + iv(x, y)$ ist approximativ stetig nach x oder y, wenn u und v es sind.“

Aus den beiden vorherigen Sätzen läßt sich nun folgern:

Hauptsatz B. *$h(z)$ sei eine komplexe Funktion im (beschränkten) Gebiete G der Ebene, deren Real- und Imaginärteil für sich in jedem in G liegenden, abgeschlossenen Bereiche B summierbar und auf jeder achsenparallelen Strecke in B auch linear summierbar sind.*

Für die Intervallfunktion $\Phi(J) = \int_J h(z)\,dz$ soll folgendes gelten:

1. *$\Phi(J)$ ist stetig* in jedem abgeschlossenen Bereich in G,* 2. *bei Gebrauch von Quadraten J kann $\limsup_{m(J)=0} \frac{|\Phi(J)|}{m(J)}$ nur in einer abzählbaren Menge in G unendlich werden,* 3. *fast überall in G ist $D\Phi = 0$.*[22])

Dann existiert eine in G analytische Funktion, welche fast überall in G mit $h(z)$ übereinstimmt.

$h(z)$ ist mit dieser analytischen Funktion identisch in G, wenn:

a) *$h(z)$ stetig ist nach x oder nach y in G, oder:*

---

a) und b) gegebenen Fällen hat man völlige Übereinstimmung. Im Falle b) folgt dies durch Anwendung des Satzes: „Eine in (a, b) approximativ stetige, beschränkte Funktion ist in allen Punkten von (a, b) die Ableitung ihres unbestimmten Integrals“ (siehe Denjoy, Bulletin Soc. Math. 1915, S. 172 u. 173).

[22]) Siehe Fußnote [20]). — Nimmt man statt 3. als Bedingung: „$\liminf_{m(J)=0} \frac{|\Phi(J)|}{m(J)} = 0$ *fast überall in G, bei Gebrauch von Quadraten J*“, so ist ein Spezialfall des obigen Satzes der folgende, Looman-Wolffsche Satz: *Im beschränkten Gebiete G ist $f(z) = u(x, y) + iv(x, y)$ analytisch regulär, falls:* 1. *u und v in G (zwei-dim.) stetig sind nach x und y,* 2. $\limsup_{m(J)=0} \frac{|\Psi(J)|}{m(J)}$, *wobei im Quadrate J: $\Psi(J) = \int_J f(z)\,dz$ ist, endlich ist in den Punkten von G vielleicht mit Ausnahme einer abzählbaren Menge,* 3. $\liminf_{m(J)=0} \frac{|\Psi(J)|}{m(J)} = 0$ *ist (bei Gebrauch von Quadraten) fast überall in G.* Siehe Nieuw Archief (Amsterdam) (2) **14**, 3 (1924), Looman und **14**, 4 (1925), Wolff.

b) *$h(z)$ auf allen in G liegenden, abgeschlossenen Strecken, welche einer fest gewählten Achse parallel laufen, beschränkt ist und approximativ stetig als Funktion der zugehörigen Veränderlichen*[23])[23a]).

Man hat zum Beweise nur die beiden ersten Sätze dieses Paragraphen anzuwenden in jedem Bereiche B, der nebst seinem Rande in G liegt.

§ 10.

Der Satz von § 7 läßt sich auf komplexes Gebiet übertragen. Das Resultat ist die folgende

Verallgemeinerung eines Satzes von Lichtenstein: *In einem (beschränkten) Gebiete G der Ebene ist $f(z) = u(x, y) + i\,v(x, y)$ analytisch regulär falls u und v diesen Bedingungen genügen:*

1. *sie sind in jedem in G liegenden, abgeschlossenen Bereiche beschränkt und in G in bezug auf jede einzelne Variable stetig;*

2. *„mit Ausnahme einer höchstens abzählbaren Menge E" existiert für jeden Punkt von G eine bestimmte Umgebung folgender Art:*

$$\frac{1}{h}[i\,\delta_x f(z) - \delta_y(z)]$$

ist für alle Werte z aus dieser Umgebung, „jedoch bei Vernachlässigung einer Menge vom Maße Null", und alle Werte $h > 0$ dem absoluten Betrage nach unterhalb einer endlichen Schranke;

3. *fast überall in G ist*

$$\lim_{h=0} \frac{1}{h}[i\,\delta_x f(z) - \delta_y f(z)] = 0, \quad h > 0.$$

Hierbei ist

$$\delta_x f(z) = f(x + h + iy) - f(x + iy)$$

und

$$\delta_y f(z) = f[x + i(y + h)] - f(x + iy).$$[24])

---

[23]) Wir haben Satz B (in seiner ersten Hälfte) in dieser etwas weitläufigen Formulierung gegeben und nicht in der einfacheren Form, welche Satz A gegeben wurde, damit deutlicher hervortritt, daß er den Looman-Wolffschen Satz umfaßt (wenn man in Satz B statt 3. die in Fußnote [20]) zuletzt genannte Bedingung einsetzt). Denn aus Stetigkeit von u und v in G nach x und y (siehe Fußnote [20])) folgt ohne weiteres nur Summierbarkeit von u und v, und Stetigkeit* von $\Psi(J)$ in jedem in G liegenden, abgeschlossenen Bereiche (vgl. § 6, letzten Teil).

[23a]) Bemerkung bei der Korrektur. Es läßt sich im Texte hinzufügen: *oder* c) *wenn zu jedem Punkte z von G eine Menge existiert, auf einer der Parallelen zur x- und y-Achse durch z, welche in z eine positive obere Dichte hat und auf der $h(z)$ stetig ist in z.*

[24]) Vgl. Lichtenstein, loc. cit. [11]) Satz IV.

Der Beweis kann auf zweierlei Art geführt werden. Erstens kann man zeigen, daß wenn

$$\Phi(J) = \int_J f(z)\,dz = \int_J u\,dx - v\,dy + i\int_J v\,dx + u\,dy = \Phi_1(J) + i\,\Phi_2(J)$$

ist, die Intervallfunktionen Φ_1 und Φ_2 den Bedingungen des Satzes von § 7 genügen und somit in G identisch Null sind. Dann sind in G auch Φ und $D\Phi$ identisch Null und daraus folgt nach dem Satze von Rademacher (§ 9) [oder nach Hauptsatz B], daß $f(z)$ in G analytisch regulär ist.

Zweitens ist es möglich, der in § 7 gegebenen Beweismethode folgend, ohne $f(z)$ und $\Phi(J)$ in reelle und imaginäre Bestandteile zu trennen, zu zeigen, daß Φ stetig* ist in jedem abgeschlossenen Intervall I in G und daß $D\Phi = 0$ ist in den Punkten von $(G - E)$. Darauf liefert Anwendung des Hauptsatzes B das gesuchte Resultat.

Korollar, Verallgemeinerung eines Satzes von Pompéiu (-Montel): *$f(z) = u(x, y) + iv(x, y)$ ist im (beschränkten) Gebiete G analytisch regulär unter den Bedingungen:*

1. *u und v sind in jedem in G liegenden, abgeschlossenen Bereiche beschränkt und in bezug auf jede einzelne Variable stetig;*

2. *„mit Ausnahme einer höchstens abzählbaren Menge E“ existiert für jeden Punkt P von G eine bestimmte Umgebung folgender Art: Die oberen und unteren Derivierten von u und v nach x und y sind beschränkt in dieser Umgebung „bei Vernachlässigung einer abzählbaren Menge von Punkten der Umgebung[25])“;*

3. *fast überall in G ist:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{und} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.\ ^{26})$$

§ 11.

Schließlich erhalten wir die beiden folgenden, etwas verallgemeinerten Sätze von Rademacher:

I. *Die komplexe Funktion $f(z) = u(x, y) + iv(x, y)$ ist analytisch regulär in G unter den Bedingungen:*

1. *in jedem abgeschlossenen Bereiche in G ist $f(z)$ flächenhaft summierbar und in x und in y für sich linear totalstetig;*

---

25) Siehe Fußnote 19).

26) Vgl. für die Herleitung aus dem ersten Satze dieses Paragraphen das Ende von § 7.

2. *fast überall genügen die partiellen Ableitungen von u und v den Cauchy-Riemannschen Differentialgleichungen auf derjenigen Menge in G, auf der alle vier existieren;*

3. *sie sind in einer Umgebung eines jeden Punktes z in G „mit Ausnahme einer abzählbaren Menge E" summierbar auf der Teilmenge dieser Umgebung, auf der alle vier existieren;*

4. $\Phi(J) = \int_J f(z)\,dz$ *ist stetig* in den Punkten der unter 3. genannten Menge E.*[27])

Daß die partiellen Ableitungen von u und v flächenhaft meßbar sind, folgt aus 1. Aus der Totalstetigkeit von u nach y folgt, daß $\frac{\partial u}{\partial y}$ in G fast überall auf jeder Parallelen zur y-Achse existiert[28]) und dadurch auch fast überall in G. Das letztere gilt auch von den drei übrigen Ableitungen. Dadurch existieren, nach 2, die C.-R. Diff.-Gl. fast überall in G.

Nehmen wir $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$ und $\frac{\partial v}{\partial y}$ gleich Null dort, wo sie nicht existieren. Dann ist in einer bestimmten Umgebung eines jeden Punktes von $(G - E)$ für ein willkürliches Intervall J $[x_1 \leqq x \leqq x_2;\ y_1 \leqq y \leqq y_2]$ (nach 3. und 1.):

$$\iint_J \frac{\partial u}{\partial y}\,dx\,dy = \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} \frac{\partial u}{\partial y}\,dy = \int_{x_1}^{x_2} [u(x, y_2) - u(x, y_1)]\,dx = -\int_J u\,dx\ ^{28}).$$

Analoge Identitäten existieren für die andern Ableitungen. Sie liefern zusammen:

$$\int_J f(z)\,dz = \int_J u\,dx - v\,dy + i\int_J v\,dx + u\,dy$$

$$= -\iint_J \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) dx\,dy + i\iint_J \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy = 0.$$

Auf $(G - E)$ ist dadurch $D\Phi = 0$.

Weiter ist, wie in § 8, zu beweisen, daß Φ stetig* ist in jedem abgeschlossenen J aus dem Gebiete G.

Nach Hauptsatz B ist nun $f(z)$ analytisch regulär in G.

II. *Im (beschränkten) Gebiete G existieren zwei abzählbare Mengen E_1 und E_2, so daß für die in G definierte Funktion $f(z) = u(x, y) + iv(x, y)$ gilt:*

1. $f(z)$ *ist in G linear stetig nach x und nach y, und auf jedem in G liegenden abgeschlossenen Bereiche summierbar;*

---

[27]) Vgl. Rademacher, loc. cit. [21]) Satz II', und Pollard, loc. cit. [8]) § 9. — Ein entsprechender Satz existiert für den reellen Fall.

[28]) Siehe z. B. Carathéodory, Reelle Funktionen, § 485, Satz 2.

2. *in jedem Punkte von* $(G - E_1)$ *sind die partiellen, oberen und unteren Derivierten von u und v nach x und y endlich;*

3. *für jeden Punkt von* $(G - E_2)$ *existiert eine Umgebung U, in der alle partiellen Ableitungen von u und v auf der Teilmenge von U summierbar sind, auf der alle existieren;*

4. *die partiellen Ableitungen genügen den C.-R. Differentialgleichungen fast überall auf derjenigen Menge in G, auf der alle vier existieren;*

5. $\Phi(J) = \int_J f(z)\,dz$ *ist stetig* in den Punkten von* E_2.

Dann ist $f(z)$ *analytisch regulär in G*[29]).

Das Beweisverfahren des § 8 ergibt, daß auf $(G - E_2)$: $D\Phi = 0$ ist und daß Φ stetig* ist in jedem abgeschlossenen, in G liegenden Intervall.

Damit sind die Bedingungen des Hauptsatzes B in G erfüllt.

§ 12.

In den Schlußparagraphen werden zwei Sätze C und D abgeleitet, welche neuere Bedingungen enthalten, unter denen sich der Cauchysche Integralsatz im reellen Fall bzw. die analytische Regularität von $f(z)$ im komplexen Fall zeigen läßt. Voran gehen einige Hilfssätze.

$f(x, y)$ sei eine im beschränkten Gebiete G definierte, reelle Funktion, für die definiert wird:

$$w_f(x, y; \varrho) = \text{obere Grenze von } \left|\frac{f(x+h, y+k) - f(x, y)}{\sqrt{h^2+k^2}}\right|$$

$$\text{für } 0 < h^2 + k^2 \leqq \varrho^2$$

und

$$L_f(x, y) = \lim_{\varrho=0} w_f(x, y; \varrho).$$

Ist $f(x, y)$ *linear stetig nach x und nach y, so läßt sich zeigen, daß* $L_f(x, y)$ *flächenhaft meßbar ist.* Wir betrachten dazu zwei abzählbare Folgen von positiven und negativen Zahlen $h_1, h_2, \ldots$ und $k_1, k_2, \ldots$, welche im Intervall $(-\varrho, +\varrho)$ überall dicht liegen (ϱ positiv). Es sei

$$W_f(x, y; \varrho) = \text{obere Grenze von } \left|\frac{f(x+h_i, y+k_j) - f(x, y)}{\sqrt{h_i^2+k_j^2}}\right|$$

$$\text{für } 0 < h_i^2 + k_j^2 \leqq \varrho^2.$$

Dann läßt sich aus der linearen Stetigkeit von $f(x, y)$ nach x und nach y folgern, daß $w_f(x, y; \varrho)$ und $W_f(x, y; \varrho)$ identisch sind in jedem Punkte eines abgeschlossenen Bereiches B in G. Wir wählen ϱ so klein, daß alle Punkte im Innern von oder auf dem Kreise, um einen willkürlichen Punkt

---

[29]) Vgl. Rademacher, loc. cit. [21]) Satz III.

des abgeschlossenen B mit dem Radius ϱ beschrieben, in G liegen. In B ist $W_f(x, y; \varrho)$ Grenzfunktion von nach x und nach y stetigen Funktionen $\varphi_n(x, y; \varrho)$, welche man wie folgt erhält. Man ordnet die abzählbare Menge von Funktionen:

$$\left|\frac{f(x+h_i, y+k_j)-f(x, y)}{\sqrt{h_i^2+k_j^2}}\right| \quad \text{für} \quad 0 < h_i^2 + k_j^2 \leqq \varrho^2$$

in eine abzählbare Reihe und nimmt dann $\varphi_n(x, y; \varrho) =$ Maximum in (x, y) von den n ersten Funktionen der Reihe. Die Funktionen $\varphi_n(x, y; \varrho)$ sind in B linear stetig nach x und nach y. Daraus folgt, daß sie auch flächenhaft meßbar sind in B[10]); somit auch $\lim\limits_{n=\infty} \varphi_n(x, y; \varrho) = W_f(x, y; \varrho)$ $= w_f(x, y; \varrho)$ und dadurch wieder $\lim\limits_{\varrho=0} w_f(x, y; \varrho) = L_f(x, y)$. Da B willkürlich in G liegt, ist $L_f(x, y)$ flächenhaft meßbar in G.

Wir erinnern an folgende Definition: „Die in einer Umgebung von (x, y) definierte Funktion $f(x, y)$ heißt an dieser Stelle vollständig differenzierbar, wenn

$$f(x+h, y+k) - f(x, y) = \frac{\partial f(x, y)}{\partial x}\cdot h + \frac{\partial f(x, y)}{\partial y}\cdot k + \varepsilon_1(h, k)\cdot h + \varepsilon_2(h, k)\cdot k$$

ist, wobei ε_1 und ε_2 nach Null konvergieren sollen in allen Fällen, daß h und k sich gleichzeitig der Null nähern."

Nun läßt sich zeigen:

Hilfssatz 1. *Die notwendige und hinreichende Bedingung dafür, daß eine in G nach x und nach y linearstetige Funktion $f(x, y)$ fast überall in G total differenzierbar sei, besteht darin, daß $L_f(x, y)$ fast überall in G endlich sei.*

Die Notwendigkeit der Bedingung ist einfach zu beweisen; in jedem Punkte (x, y), in welchem $f(x, y)$ total differenzierbar ist, ist die auf Seite 149 definierte Funktion $w_f(x, y; \varrho)$ beschränkt für alle positiven ϱ unterhalb einer gewissen Schranke und ist somit $L_f(x, y) = \lim\limits_{\varrho=0} w_f(x, y; \varrho)$ endlich.

Nehmen wir, umgekehrt, die Endlichkeit von $L_f(x, y)$ auf einer Menge K in G an mit $m(K) = m(G)$. Da f stetig ist nach x und nach y, ist sie flächenhaft meßbar[16]) und dadurch sind dies auch die oberen und unteren Derivierten von f nach x oder y[16]). Die Ableitungen nach x und y existieren nun auf meßbaren Mengen in G.

In den Punkten von K ist $f(x, y)$ (zweidimensional) stetig. Nach dem Satze von Fubini gibt es dadurch eine Menge E_x auf der x-Achse vom Maße Null, so daß für jeden Wert x, welcher nicht zu E_x gehört und für den die Parallele L_x zur y-Achse Punkte von G enthält, $f(x, y)$ fast überall linear stetig ist auf dem Durchschnitte $(L_x \cdot G)$. Sie ist also auch linear meßbar auf $(L_x \cdot G)$. Da die oberen und unteren Derivierten

nach y auf K endlich sind, existiert eine E_x enthaltende Menge E'_x vom Maße Null, so daß für jeden Punkt x, der nicht zu E'_x gehört und dessen Parallele L_x zur y-Achse Punkte von G enthält, $f(x, y)$ fast überall auf $(L_x \cdot G)$ endliche obere und untere Derivierten besitzt. Nach einem Satze von G. C. Young[17]) besitzt dadurch $f(x, y)$ fast überall auf $(L_x \cdot G)$ eine endliche Ableitung $\frac{\partial f}{\partial y}$, und somit auch fast überall in G.

Auf dieselbe Weise zeigt man, daß $\frac{\partial f}{\partial y}$ fast überall in G existiert.

Von hier an fährt der Beweis fort wie in den §§ 2 und 3 einer Arbeit von Rademacher: „Über partielle und totale Differenzierbarkeit usw.", Math. Annalen Bd. 79 (1919). Dort enthält Satz 1 das Resultat: „Ist $f(x, y)$ eine in dem beschränkten Gebiete G definierte Funktion, für welche $L_f(x, y)$ in G endlich und summierbar ist, so ist $f(x, y)$ fast überall in G total differenzierbar[30])." Auch die weniger fordernden Bedingungen des Hilfssatzes 1 führen zu derselben Eigenschaft für die Funktion $f(x, y)$.

§ 13.

Hilfssatz 2. *Ist in der Umgebung eines Punktes (ξ, η) $p(x, y)$ linear summierbar nach x und $q(x, y)$ nach y und sind beide Funktionen in (ξ, η) total differenzierbar, so ist für die Intervallfunktion $\Phi(J) = \int_J p\,dx + q\,dy$ in (ξ, η) die (in § 1 definierte) Ableitung* $= \frac{\partial q(\xi, \eta)}{\partial x} - \frac{\partial p(\xi, \eta)}{\partial y}$.

Es sei (ξ, η) innerer Punkt oder Randpunkt eines Quadrates J $[\xi - t_1 \leqq x \leqq \xi + t_2;\ \eta - u_1 \leqq y \leqq \eta + u_2]$. Dann ist:

$$\int_J p\,dx = \int_{-t_1}^{+t_2} [p(\xi + t, \eta - u_1) - p(\xi + t, \eta + u_2)]\,dt$$

$$= \int_{-t_1}^{+t_2} [\{p(\xi + t, \eta - u_1) - p(\xi, \eta)\} - \{p(\xi + t, \eta + u_2) - p(\xi, \eta)\}]\,dt$$

$$= \int_{-t_1}^{+t_2} \Big[-\frac{\partial p(\xi, \eta)}{\partial y} \cdot u_1 + \varepsilon_1(t, u_1) \cdot t - \varepsilon_2(t, u_1) \cdot u_1$$
$$- \frac{\partial p(\xi, \eta)}{\partial y} \cdot u_2 - \varepsilon_3(t, u_2) \cdot t - \varepsilon_4(t, u_2) \cdot u_2\Big]\,dt$$

$$= -\frac{\partial p(\xi, \eta)}{\partial y} \cdot (u_1 + u_2)(t_1 + t_2) + \int_{-t_1}^{+t_2} [\varepsilon_1(t, u_1) - \varepsilon_3(t, u_2)] \cdot t\,dt$$
$$+ \int_{-t_1}^{+t_2} \varepsilon_5(t, u_1, u_2) \cdot [u_1 + u_2]\,dt.$$

---

30) Eine Verallgemeinerung, welche von Stepanoff gefunden wurde nach dem Verfahren, das im Texte zu Hilfssatz 1 führte, findet man in den Math. Annalen **90** (1923).

Hierbei konvergieren die Funktionen $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_5$ nach Null, wenn sich das Quadrat in (ξ, η) zusammenzieht. Es existiert eine positive Funktion $\varepsilon_6(t_1, t_2, u_1, u_2)$, welche nach Null konvergiert, wenn t_1, t_2, u_1 und u_2 sich der Null nähern und welche die Eigenschaft besitzt, daß der absolute Wert der Summe der beiden letzten Integrale gleich

$$\varepsilon_6(t_1, t_2, u_1, u_2) \cdot \int_{-t_1}^{+t_2} [|t| + (u_1 + u_2)]\, d|t| < \varepsilon_6 \cdot \frac{3}{2} m(J)$$

ist.

Durchläuft J eine willkürliche Reihe den Punkt (ξ, η) als inneren oder Randpunkt enthaltender und sich in (ξ, η) zusammenziehender Quadrate, so folgt:

$$\lim_{m(J)=0} \frac{\int_J p\, dx}{m(J)} = -\frac{\partial p(\xi, \eta)}{\partial y}.$$

Dann findet man auch auf gleiche Weise:

$$\lim_{m(J)=0} \frac{\int_J q\, dy}{m(J)} = \frac{\partial q(\xi, \eta)}{\partial x},$$

womit der Beweis fertig ist [31]).

§ 14.

Hilfssatz 3. *Die Funktionen $p(x, y)$ und $q(x, y)$ sind in der Umgebung von (ξ, η) linear summierbar nach x bzw. nach y. Die (wie in § 12 definierten) Funktionen $L_p(x, y)$ und $L_q(x, y)$ sind in (ξ, η) kleiner als ein positives M. Dann läßt sich zeigen, daß für die Intervallfunktion $\Phi(J) = \int_J p\, dx + q\, dy$ die obere und untere Derivierten in (ξ, η) ihrem absoluten Werte nach kleiner sind als $4 M \sqrt{2}$.*

---

[31]) Existiert eine Funktion $f(x, y)$ in der Umgebung von (ξ, η), so daß in dieser Umgebung die Ableitungen $\frac{\partial f}{\partial x} = p(x, y)$ und $\frac{\partial f}{\partial y} = q(x, y)$ endlich sind, und sind p und q in (ξ, η) total differenzierbar, so sind in der Umgebung von (ξ, η) p nach x und q nach y summierbar. Nach Hilfssatz 2 ist also $D_{(\xi, \eta)} \Phi(J) = \frac{\partial q(\xi, \eta)}{\partial x} - \frac{\partial p(\xi, \eta)}{\partial y}$, wobei $\Phi(J) = \int_J p\, dx + q\, dy$ ist. Weiter ist einfach zu sehen, daß $\Phi(J) \equiv 0$ ist in der Umgebung von (ξ, η). Somit ist auch $D_{(\xi, \eta)} \Phi(J) = 0$ oder $\frac{\partial p(\xi, \eta)}{\partial y} = \frac{\partial q(\xi, \eta)}{\partial x}$. Damit ist ein neuer Beweis des folgenden Satzes von W. H. Young [Proc. London Math. Soc. (2) 7 (1909)] geliefert: *Besitzt eine Funktion $f(x, y)$ in der Umgebung eines Punktes (ξ, η) partielle Ableitungen $p(x, y)$ nach x und $q(x, y)$ nach y und sind diese Ableitungen in (ξ, η) total differenzierbar, so ist in (ξ, η): $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$.*

Da $L_p(x, y)$ kleiner ist als M, läßt sich ein positives ϱ angeben, so daß für jeden Punkt $(\xi + t, \eta + u)$ mit $0 < t^2 + u^2 \leqq \varrho^2$ gilt:

$$\left| \frac{p(\xi + t, \eta + u) - p(\xi, \eta)}{\sqrt{t^2 + u^2}} \right| < M.$$

Es sei J ein Quadrat, das (ξ, η) im Innern oder auf dem Rande enthält und das ganz im Innern des Kreises $K(\varrho)$, vom Radius ϱ um (ξ, η), liegt. Genügen die Koordinaten den Ungleichungen:

$$\xi - t_1 \leqq x \leqq \xi + t_2;\ \eta - u_1 \leqq y \leqq \eta + u_2,$$

so ist:

$$\int_J p\,dx = \int_{-t_1}^{+t_2} [\{p(\xi + t, \eta - u_1) - p(\xi, \eta)\} - \{p(\xi + t, \eta + u_2) - p(\xi, \eta)\}]\,dt$$

$$= M \int_{-t_1}^{+t_2} [\Theta_1(t, u_1) \cdot \sqrt{t^2 + u_1^2} - \Theta_2(t, u_2) \cdot \sqrt{t^2 + u_2^2}]\,dt,$$

wobei die beiden Funktionen Θ_1 und Θ_2, absolut genommen, kleiner als 1 sind. Hieraus folgt leicht:

$$\left| \int_J p\,dx \right| < M \times 2(t_1 + t_2)\sqrt{2} \int_{-t_1}^{+t_2} dt = M \times 2\sqrt{2} \times m(J).$$

Da übereinstimmende Betrachtungen gelten für $q(x, y)$, ist der Beweis leicht zu beendigen.

§ 15.

Nun folgt der

Satz C: $\Phi(J) = \int_J p\,dx + q\,dy$ *ist im (beschränkten) Gebiete* $G \equiv 0$ *unter den Bedingungen:*

1. *p und q sind beschränkt in jedem in G liegenden, abgeschlossenen Bereiche und in bezug auf jede einzelne Variable stetig,*

2. *$L_p(x, y)$ und $L_q(x, y)$ können nur in einer abzählbaren Menge E unendlich werden.*

3. *fast überall in G ist* $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$. [32])

Wie in § 7, folgt aus 1., daß Φ stetig* ist in jedem in G liegenden, abgeschlossenen Intervall I.

Aus 1. und 2. folgt, nach Hilfssatz 1, daß p und q fast überall in G total differenzierbar sind. Dieses und 3. liefert, nach Hilfssatz 2, daß $D\Phi$ fast überall Null ist in G.

---

[32]) Diese Bedingung hätte man auch schreiben können: „fast überall ist $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$ auf derjenigen Menge in G, auf der beide existieren". Vgl. S. 141.

Nach 2. und Hilfssatz 3 sind in allen Punkten von G, die Menge E ausgenommen, obere und untere Derivierten von $\Phi(J)$ endlich.

Im offenen Intervall I sind die Bedingungen des Hauptsatzes A erfüllt. Daraus folgt: $\Phi \equiv 0$ in I-, also auch in G.

Nimmt man statt 1. *zwei-dim. Stetigkeit von p und q in G an, so verschwindet das Linienintegral auch über jede einfache, geschlossene, rektifizierbare Kurve in G* [8]).

§ 16.

Der übereinstimmende Satz bei komplexen Funktionen lautet:

Satz D: *Die komplexe Funktion $f(z) = u(x, y) + iv(x, y)$ ist analytisch regulär im (beschränkten) Gebiete G, falls:*

1. *$f(z)$ in jedem in G liegenden, abgeschlossenen Bereiche B beschränkt ist und in bezug auf jede einzelne Variable stetig,*

2. $\limsup\limits_{h=0} \left|\frac{f(z+h)-f(z)}{h}\right|$ *nur auf einer abzählbaren Menge E in G unendlich werden kann.*

3. *fast überall in G die Cauchy-Riemannschen Differentialgleichungen gelten.*

Aus 1. ist abzuleiten, daß $\Phi(J) = \int\limits_J f(z)\,dz$ in jedem abgeschlossenen Intervall I in G stetig* ist [33]).

Aus 2. folgt, daß $L_u(x, y)$ und $L_v(x, y)$ endlich sind in $(G - E)$. Dies und 1. liefert, nach Hilfssatz 1, daß u und v fast überall in G total differenzierbar sind. Es sei $\Phi(J) = \Phi_1(J) + i\,\Phi_2(J)$. Dann folgt mit 3., nach Hilfssatz 2, daß $D\,\Phi = D\,\Phi_1 + i\,D\,\Phi_2$ fast überall in G Null ist.

Nach 2. und Hilfssatz 3 können die oberen und unteren Derivierten von Φ_1 und Φ_2 nur auf E unendlich werden.

Schließlich folgt aus 1., daß $f(z)$ summierbar ist in jedem abgeschlossenen Bereiche B in G [10]).

Die Bedingungen des Hauptsatzes B sind also erfüllt und $f(z)$ ist analytisch regulär in G [34]).

---

[33]) Vgl. §§ 7 und 10.

[34]) Nach Montel [Comptes Rendus **156** (1913), p. 1820–1822] ist $f(z)$ (unter Annahme zwei-dimensionaler Stetigkeit) auch analytisch regulär, wenn man *statt* 2. *Endlichkeit der partiellen Ableitungen* annimmt. Ein Beweis wurde von ihm jedoch nicht gegeben. Ein Beweis von Borger [Bull. Am. Math. Soc. 27 (1921); er nimmt außerdem die C.-R.-Diff.-Gleichungen in „allen" Punkten von G an] ist, wie Looman [Gött. Nachr. (1923)] angibt, falsch. Aber auch dieser liefert [loc. cit.] keinen richtigen Beweis.

Er wünscht erstens zu zeigen: „In einem beschränkten Gebiete G ist für jedes achsenparallele Rechteck r: $\int\limits_r p\,dx + q\,dy = 0$, wenn p und q (zwei-dim.) stetig sind

Fortsetzung der Fußnote [34]) auf nächster Seite.)

Der bekannte Goursatsche Satz ist ein Korollar des obigen Satzes. Auch der verallgemeinerte Satz von Pompéiu(-Montel); denn aus Beschränktheit der partiellen Derivierten nach x und y in einem abgeschlossenen Intervall J folgt Beschränktheit der zugehörigen Differenzenquotienten[14]),

---

in G, $\frac{\partial p}{\partial y}$ und $\frac{\partial q}{\partial x}$ endlich sind in G und fast überall in G: $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$ ist." Daraus würde dann leicht das verlangte Resultat folgen.

Zum Beweise dieses Satzes wird eine bestimmte Menge M definiert, welche die Eigenschaften besitzt, perfekt und nirgends dicht zu sein. Darauf wird gezeigt, daß in jedem abgeschlossenen Intervall r in G, das in seinem Innern Punkte von M enthält, ein abgeschlossenes Intervall r_1 liegt, das auch Punkte von M im Innern enthält und folgende Eigenschaft hat. Gehört (x, y) zu $(r_1 \cdot M) = \sigma(r_1)$ und gehört $(x, y+h)$ zu r_1, so liegt $\left|\frac{p(x, y+h) - p(x, y)}{h}\right|$ unterhalb einer endlichen Schranke A. Das abgeschlossene Intervall r_2 $[a_2 \leqq x \leqq b_2;\ c_2 \leqq y \leqq d_2]$ liege im Innern von r_1. Enthält die von r_2 auf einer Parallelen L_x zur y-Achse ausgeschnittene, abgeschlossene Strecke Punkte von $(r_2 \cdot M) = \sigma(r_2)$, so wird sie in zwei Mengen geteilt, und zwar in die perfekte Menge $(L_x \cdot \sigma(r_2))$ und die Menge der offenen, komplementären Intervalle $(\gamma_i \delta_i)$. Nun läßt sich zeigen, daß

$$p(x, d_2) - p(x, c_2) = \Sigma[p(x, \delta_i) - p(x, \gamma_i)] + \int\limits_{x, \sigma(r_2)} \frac{\partial p}{\partial y} \partial y,$$

wobei dann gilt:

$$(1) \qquad |\Sigma[p(x, \delta_i) - p(x, \gamma_i)]| < A \cdot \Sigma(\delta_i - \gamma_i).$$

Darauf wird eine Reihe von abgeschlossenen Quadraten r_3 betrachtet, welche einen Punkt (x, y) von $\sigma(r_1)$ enthalten und sich auf (x, y) zusammenziehen. Die Parallelen zur y-Achse, welche ein r_3 überschneiden, enthalten Punkte von $\sigma(r_3)$ oder nicht; die zugehörigen x-Werte der ersten Gruppe von Parallelen bilden eine Menge $E_1(x)$ und die der zweiten Gruppe eine Menge $E_2(x)$. Nun ist:

$$\int\limits_{r_3} p\, dx = -\int\limits_{a_3}^{b_3} dx \int\limits_{x, \sigma(r_3)} \frac{\partial p}{\partial y} dy - \int\limits_{E_1(x)} dx \cdot \Sigma[p(x, \delta_i) - p(x, \gamma_i)] - \int\limits_{E_2(x)} dx \cdot [p(x, d_3) - p(x, c_3)].$$

Die beiden letzten Integrale werden in der Loomanschen Arbeit zusammengenommen, und es wird aus (1) irrtümlich gefolgert, daß der absolute Wert ihrer Summe $< A \cdot \mu[r_3 - \sigma(r_3)]$, wobei μ das Flächenmaß andeutet. Aus (1) folgt zwar für das Integral über $E_1(x)$, daß sein absoluter Wert kleiner als $A \cdot \{\mu[r_3 - \sigma(r_3)] - (d_3 - c_3) \cdot m E_2(x)\}$ ist, wobei μ Flächenmaß und m lineares Maß andeutet. *Nun müßte jedenfalls noch bewiesen werden:*

$$(2) \qquad \left|\int\limits_{E_2(x)} dx \cdot [p(x, d_3) - p(x, c_3)]\right| < A \cdot (d_3 - c_3) \cdot m E_2(x),$$

damit die von Looman gegebene Abschätzung für die Summe beider Integrale richtig wäre. (2) wäre bewiesen, wenn $\left|\frac{p(x, d_3) - p(x, c_3)}{d_3 - c_3}\right| < A$ wäre auf $E_2(x)$ für *alle* Quadrate r_3 der sich auf (x, y) zusammenziehenden Reihe. Aber das ist ohne weiteres nicht einzusehen. Dadurch liefern natürlich die weiteren Schlüsse der zitierten Arbeit nicht den gesuchten Beweis.

dadurch auch von $\frac{u(x+h,\, y+k)-u(x,y)}{\sqrt{h^2+k^2}}$ und $\frac{v(x+h,\, y+k)-v(x,y)}{\sqrt{h^2+k^2}}$, und dadurch wieder von $\limsup\limits_{h=0} \left|\frac{f(z+h)-f(z)}{h}\right|$ in J.

§ 17.

Daß die Bedingung der Beschränktheit von $f(z)$ in jedem in G liegenden, abgeschlossenen Bereiche nicht fortgelassen werden darf, zeigt das Beispiel: $f(z) = e^{-\frac{1}{z^4}}$. Diese Funktion genügt in der Umgebung von $[x=0,\ y=0]$ allen Bedingungen des Satzes D, ausgenommen die der Beschränktheit. Sie ist auch nicht analytisch regulär in dieser Umgebung[35]).

[35]) Das Beispiel entlehnen wir Looman, loc. cit. [34]).

(Eingegangen am 15. 11. 1928.)

On the polynomial and trigonometric approximation of measurable bounded functions on a finite interval.

Von

J. Shohat in Ann Arbor (Mich., U. S. A.).

Introduction.

Let $f(x)$ be a continuous function defined on a finite interval (a, b). The important role of Weierstraß' theorem dealing with polynomial (or trigonometric) approximation of such functions is well known. With this theorem is closely connected the important notion, due to Tchebycheff, of the "polynomial of the best approximation" to $f(x)$ on (a, b), of degree $\leqq n$ — it will be denoted here by $\Pi_n(x)$ —, for which the "deviation" from $f(x)$ — the so called "best approximation" $E_n(f) = \max|f(x) - \Pi_n(x)|$ on (a, b) — is the smallest possible, compared with any other polynomial of degree $\leqq n$. However, the actual construction of $\Pi_n(x)$ is attainable in a very limited number of cases only.

It is, therefore, of interest to give *for any continuous function $f(x)$, defined on a finite interval (a, b), a sequence of polynomials of degree $n = 1, 2, \ldots$ which, as $n \to \infty$, converges uniformly to $f(x)$ throughout the whole interval (a, b), and yields, for n very large, an approximation of the same order as that of the best approximation.*

This was the original object of this paper attained by considering the minimum of the integral $\int_a^b p(x)|f(x) - P_{nm}(x)|^m dx$ $(m \geqq 1)$, for $m, n \to \infty$, where $p(x)$ $(\geqq 0)$ and $f(x)$ are properly defined on (a, b) and $P_{nm}(x)$ is the required minimizing polynomial of degree $\leqq n$. The author wishes to acknowledge this part as an outgrowth of a correspondence with Professor Paul Lévy of the École Polytechnique in Paris.

In the course of the said investigation it was found possible to *extend in two ways the notion of the polynomial of the best approximation* (*in*

the above sense of Tchebycheff) to the more general class of measurable bounded functions, making use of "measurable bounds" introduced for such functions by C. N. Haskins.

We consider in the present paper the above minimum for all possible cases: 1. m is fixed, $n \to \infty$; 2. n is fixed, $m \to \infty$; 3. $m, n \to \infty$. Thus, our results supplement and generalize those previously given by G. Pólya, D. Jackson and the writer [1]).

§ 1.

In our discussion we shall make frequent use of the following inequalities:

$$(1)\qquad \left|\int_a^b f_1(x) f_2(x)\,dx\right| \leqq \left[\int_a^b |f_1(x)|^s dx\right]^{\frac{1}{s}} \left[\int_a^b |f_2(x)|^{\frac{s}{s-1}} dx\right]^{\frac{s-1}{s}} \quad (s>1)$$

$$(2)\qquad \left[\int_a^b |f_1(x)+f_2(x)|^s dx\right]^{\frac{1}{s}} \leqq \left[\int_a^b |f_1(x)|^s dx\right]^{\frac{1}{s}} + \left[\int_a^b |f_2(x)|^s dx\right]^{\frac{1}{s}}$$

$$(s \geqq 1)\ ^{2})$$

$$(3)\qquad \left[\int_a^b |f_1(x)-f_2(x)|^s dx\right]^{\frac{1}{s}} \geqq \left[\int_a^b |f_1(x)|^s dx\right]^{\frac{1}{s}} - \left[\int_a^b |f_2(x)|^s dx\right]^{\frac{1}{s}}$$

$$(4)\qquad \int_a^b |f_1(x)|\,|f_2(x)|^{s_1} dx \leqq \left[\int_a^b |f_1(x)|\,|f_2(x)|^{s_2} dx\right]^{\frac{s_1}{s_2}} \left[\int_a^b |f_1(x)|\,dx\right]^{\frac{s_2-s_1}{s_2}} \quad (s_2 \geqq s_1 > 0)$$

$$(5)\qquad |a+b|^s \leqq 2^{s-1}[|a|^s + |b|^s] \quad (s \geqq 1).$$

In (1 — 4) the existence of the right-hand integrals (integrals are taken in the sense of Lebesgue throughout this paper) implies the existence of those on the left side.

Hereafter, the following general notations will be used: $G_n(x) = \sum_{i=0}^{n} g_i x^i$ — to denote an ***arbitrary polynomial*** of degree $\leqq n$, subject in some cases to certain explicitly stated conditions; A, ε — to denote respectively a suffi-

---

[1]) a) G. Pólya, Sur un algorithme toujours convergent pour les polynomes de la meilleure approximation de Tchebycheff pour une fonction continue quelconque, Comptes Rendus **157** (1913), p. 840–843. b) D. Jackson, On the convergence of certain trigonometric and polynomial approximations, Transactions of the American Mathematical Society **22** (1921), p. 158–166. c) Idem, Note on the convergence of weighed trigonometric series, Bulletin of the American Mathematical Society **29** (1923), p. 259–263. d) J. Shohat, On the polynomial of the best approximation to a given continuous function, ibid. **31** (1925), p. 509–514.

[2]) F. Riesz, Über Systeme integrierbarer Funktionen, Math. Annalen **69** (1911), S. 449–497, S. 456.

ciently large or sufficiently small, but fixed positive quantity properly chosen in each case; $m_0, n_0, \ldots$ — to denote properly chosen sufficiently large numbers of given sets $(m), (n), \ldots$; τ, τ_{nm} — to denote properly chosen fixed positive quantities which remain finite as $m, n \to \infty$.

§ 2.

Definition. $p(x)$, defined on a given interval (a, b) — finite or infinite — will be called a *"characteristic function"*, abbreviatedly, *"c-function"*, if: α) $p(x) \geqq 0$ *in* (a, b), β) *all integrals* $\int_a^b p(x)\, x^i\, dx$ $(i = 0, 1, \ldots)$ *exist with* $\int_a^b p(x)\, dx > 0$. In case of finite (a, b) the integrability of $p(x)$ on (a, b) implies the existence of all the integrals above.

With any such c-function $p(x)$ we can construct a system of orthogonal and normal Tchebycheff polynomials

$$\varphi_n(p; x) \equiv \varphi_n(x) = a_n(p)\, x^n + \ldots \quad (n = 0, 1, 2, \ldots;\ a_n > 0) \tag{6}$$

$$\int_a^b p(x)\, \varphi_n(x)\, \varphi_m(x)\, dx = \begin{cases} 0, & m \neq n \\ 1, & m = n, \end{cases} \tag{7}$$

which enable us to solve the following problem.

Find the upper limit of $|\omega(G_n)| \equiv \left| \sum_{i=0}^{n} \alpha_i g_i \right|$ *for all polynomials* $G_n(x) = \sum_{i=0}^{n} g_i x^i$, *of degree* $\leqq n$, satisfying the inequality

$$\int_a^b p(x)\, |G_n(x)|^m\, dx \leqq M^m.$$

Here $p(x)$ *designates a c-function defined on* (a, b), $\alpha_i\ (i = 0, 1, \ldots, n)$, $M\ (> 0)$ *and* $m\ (\geqq 2)$ *are certain given constants.*

Solution. Applying (1) we get,

$$\int_a^b p(x)\, G_n^2(x)\, dx \leqq M^2 \left[\int_a^b p(x)\, dx \right]^{\frac{m-2}{m}}.$$

Thus, our problem is reduced to the case $m = 2$, and we can apply the solution previously given by the writer [8]):

$$|\omega(G_n)| \leqq M \left[\int_a^b p(x)\, dx \right]^{\frac{m-2}{2m}} \left\{ \sum_{i=0}^{n} \omega^2(\varphi_i) \right\}^{\frac{1}{2}}. \tag{8}$$

---

[8]) J. Shohat, On a general formula in the theory of Tchebycheff polynomials and its applications, Transactions of the American Mathematical Society 29 (1927), p. 569–583; p. 569–571.

Special case: $\omega(G_n) = G_n(z)$, z real arbitrary.

$$\int_a^b p(x)\,|G_n(x)|^m\,dx \leqq M^m$$

implies:

$$|G_n(z)| \leqq M K_n^{\frac{1}{2}}(z)\Big[\int_a^b p(x)\,dx\Big]^{\frac{m-2}{2m}} \tag{9}$$

$$K_n(p;x) \equiv K_n(x) \equiv \sum_{i=0}^{n} \varphi_i^2(x). \tag{10}$$

The assumption $m \geqq 2$ is sufficient for the applications below. In case $2 > m \geqq 1$, the representation of $\omega(G_n)$ as a definite integral used in deriving (8) gives,

$$|\omega(G_n)| \leqq M\Big[\int_a^b p(x)\,dx\Big]^{\frac{m-1}{m}} \sum_{i=0}^{n} |\omega(\varphi_i)|\,\Phi_i \tag{11}$$

$$(\Phi_i = \max|\varphi_i(x)| \text{ in } (a,b)).$$

Note. *The above upper limits for $|\omega(G_n)|$ do not depend on m, if M and the a_i do not,* for in (8), (9) $\int_a^b p(x)\,dx \leqq 1$ or $\leqq \Big[\int_a^b p(x)\,dx\Big]^{\frac{1}{2}}$, according to whether $\int_a^b p(x)\,dx \leqq 1$ or > 1, and similarly in (11).

§ 3.

We apply the solution of the aforesaid problem to the proof of

Theorem 1. *Given on (a,b) a c-function $p(x)$ and a measurable function $f(x)$ "of the clas $[L_p^m]$", i. e. such that $\int_a^b p(x)\,|f(x)|^m\,dx$ exists $(m \geqq 1)$. Among all polynomials of degree $\leqq n$ there exists at least one $P_{nm}(x)$ minimizing the integral $\int_a^b p(x)\,|f(x) - P_{nm}(x)|^m\,dx$. In case $m > 1$ this polynomial is unique*[4]).

[4]) This minimum problem, with more or less restricted $p(x)$ and $f(x)$, has been treated previously (loc. cit. [1])). Here the only condition imposed upon $p(x)$ is: $\int_a^b p(x)\,dx > 0$. This clearly is no restriction, since the hypothesis $\int_a^b p(x)\,dx = 0$ implies: $p(x) = 0$ almost everywhere in (a,b). Similarly, the condition of existence of $\int_a^b p(x)\,|f(x)|^m\,dx$ is imposed by the very nature of our problem, for, as it is readily seen from (1–4), the integrals $\int_a^b p(x)\,|f(x)|^m\,dx$, $\int_a^b p(x)\,|f(x) - G_n(x)|^m\,dx$ exist or do not exist simultaneously.

Proof. By (3),

$$\left[\int_a^b p(x)\,|f(x)-G_n(x)|^m dx\right]^{\frac{1}{m}} \geqq \left[\int_a^b p(x)\,|G_n(x)|^m dx\right]^{\frac{1}{m}} - \left[\int_a^b p(x)\,|f(x)|^m dx\right]^{\frac{1}{m}}.$$

The assumption $\int_a^b p(x)\,|G_n(x)|^m dx > (2A)^m$, with $A^m > \int_a^b p(x)\,|f(x)|^m dx$, implies: $\int_a^b p(x)\,|f(x)-G_n(x)|^m dx > A^m$. Therefore, we confine ourselves to polynomials $G_n(x)$ such that

$$\int_a^b p(x)\,|G_n(x)|^m dx \leqq (2A)^m.$$

The coefficients of such $G_n(x)$, for arbitrarily given n, are necessarily *bounded*, as we learn from (8), taking successively $\alpha_i = 1$ $(i = 0, 1, \ldots, n)$, $\alpha_j = 0$ $(j \neq i)$, and the existence of one at least minimizing polynomial $P_{nm}(x)$ follows.

As to the uniqueness of $P_{nm}(x)$ for $m > 1$, the proof is identical to that given in my aforesaid paper [5]).

The above proof gives incidentally the following

Corollary. *To an arbitrarily large $A > 0$ there correspond certain K_i $(i = 0, 1, \ldots, n)$ such that any one of the inequalities $|g_i| > K_i$ implies: $\int_a^b p(x)\,|f(x)-G_n(x)|^m dx > A$, where the c-function $p(x)$ and the function $f(x)$ of the class $[L_p^m]$ are arbitrary. The K_i do not depend on m, if A and M do not; $m \geqq 1$.*

We shall use the notation

$$(12)\quad I_{nm} = \int_a^b p(x)\,|f(x)-P_{nm}(x)|^m dx = \min \int_a^b p(x)\,|f(x)-G_n(x)|^m dx,$$

and we have evidently,

$$(13)\qquad I_{nm} \leqq \int_a^b p(x)\,|f(x)|^m dx.$$

§ 4.

The following two theorems, interesting by themselves, are needed for the investigation of I_{nm} and $P_{nm}(x)$.

Theorem II. 1°. *Let $p(x)$ be non-negative and integrable on (a, b) and such that $\int_E p(x)\,dx > 0$, E denoting an arbitrary measurable set of*

[5]) Loc. cit. [1]), d), p. 511. Uniqueness, in case $m = 1$, seems to require additional conditions for $p(x)$.

points in (a, b) *with* $mE > 0$. *Let the measurable function* $f(x)$, *defined on* (a, b), *be such that* $I_m = \int_a^b p(x)|f(x)|^m dx$ *exists for every* $m > 0$ [6]). *Denote by* F *the "measurable upper bound"* [7]) *of* $|f(x)|$ *in* (a, b). *Then,* $\lim_{m\to\infty} I_m^{\frac{1}{m}} = F$. 2°. *If* m *monotonically increases without bounds, then* $\lim_{m\to\infty} I_m^{\frac{1}{m}}$ *exists,* $\leqq F$, *for any* $p(x)$ *non-negative and integrable in* (a, b), *and* $\left[\frac{I_m}{\int_a^b p(x)dx}\right]^{\frac{1}{m}}$ *is also monotonically increasing* $\Big($*so is* $I_m^{\frac{1}{m}}$, *if* $\int_a^b p(x)dx \leqq I\Big)$.

Proof. 1°. The proof is somewhat similar to that given by Prof. Haskins for the case $p(x) = 1$.[8]) By the definition of "measurable upper bound", we write,

$$mE[|f(x)| \geqq F - \varepsilon] > 0, \quad mE[|f(x)| \geqq F + \varepsilon] = 0 \qquad (a \leqq x \leqq b)$$

(with obvious modifications, in case $F = +\infty$). If F be finite, let E_ε denote the set of points x in (a, b) such that $|f(x)| \geqq F - \varepsilon$. Then,

$$I_m^{\frac{1}{m}} \geqq (F-\varepsilon)\Big[\int_{E_\varepsilon} p(x)\,dx\Big]^{\frac{1}{m}}.$$

On the other hand,

$$I_m^{\frac{1}{m}} \leqq (F+\varepsilon)\Big[\int_a^b p(x)\,dx\Big]^{\frac{1}{m}}. \tag{14}$$

Therefore, since ε does not depend on m,

$$\varliminf_{m\to\infty} I_m^{\frac{1}{m}} \geqq F, \qquad \varlimsup_{m\to\infty} I_m^{\frac{1}{m}} \leqq F. \tag{15}$$

If F be $+\infty$, denote by E_A the set of points x in (a, b) such that $|f(x)| \geqq 2A$. Then,

$$I_m^{\frac{1}{m}} > 2A\Big[\int_{E_A} p(x)\,dx\Big]^{\frac{1}{m}} > A \qquad (m \geqq m_0). \tag{16}$$

The inequalities (15), (16) prove our statement.

---

[6]) This follows from its existence, say, for all sufficiently large integral m.

[7]) C. N. Haskins, On the measurable bounds and the distribution of functional values of summable functions. Transactions of the American Mathematical Society **17** (1916), p. 181–194, p. 184.

[8]) *Loc. cit.* [7]), p. 187–188.

2°. This follows from (14) and from the inequality (see (4))

$$I_{m_1}^{\frac{1}{m_1}} \leq I_{m_2}^{\frac{1}{m_2}} \Big[\int_a^b p(x)\,dx\Big]^{\frac{m_2-m_1}{m_1 m_2}} \qquad (m_2 > m_1 > 0). \tag{17}$$

Note. The condition $\int_E p(x)\,dx > 0$ $(mE > 0)$ is indispensable, as it is seen by taking $p(x) = 0$, $f(x) = 1$ for $a < \alpha \leq x \leq \beta < b$ and $p(x) = 1$, $f(x) = 0$ elsewhere in (a, b).

Theorem III. *Given on a finite interval* (a, b) *a non-negative integrable function* $p(x)$ *and a family* $\{f(x, m)\}$ *of continuous functions, where the parameter* $m (> 0)$ *takes a set of values* $\to \infty$. *Assume:* α) $\lim_{m\to\infty} f(x, m) = f(x)$ *uniformly for* $a \leq x \leq b$; β) *among the points in* (a, b), *where* $|f(x)|$ *(necessarily continuous) attains its maximum* F, *there exists at least one, say,* $x = c$, *such that* $\int_{(\delta)} p(x)\,dx > 0$, (δ) *denoting an arbitrary sub-interval of* (a, b) *containing the point* c. *Then,* $\lim_{m\to\infty} \Big[\int_a^b p(x)\,|f(x, m)|^m\,dx\Big]^{\frac{1}{m}} = F$.

Proof. First, we have

$$|f(x, m) - f(x)| < \varepsilon, \qquad |f(x, m)| < F + \varepsilon \qquad (a \leq x \leq b;\ m \geq m_0).$$

This assures the existence, for all m under consideration, of $\int_a^b p(x)\,|f(x, m)|^m\,dx$, with

$$i_m^{\frac{1}{m}} \equiv \Big[\int_a^b p(x)\,|f(x, m)|^m\,dx\Big]^{\frac{1}{m}} < (F + \varepsilon)\Big[\int_a^b p(x)\,dx\Big]^{\frac{1}{m}} \quad (m \geq m_0). \tag{18}$$

On the other hand, consider the point $x = c$ and the sub-interval (δ) with the properties given above. In virtue of the continuity of $f(x)$ and the uniform convergence of $f(x, m)$, we can fixe $m = m_1$ so large and the interval (δ) so small as to have

$$|f(x, m)| \geq F - \varepsilon \qquad (x \text{ belongs to } (\delta);\ m \geq m_1),$$

$$i_m^{\frac{1}{m}} \geq (F - \varepsilon)\Big[\int_{(\delta)} p(x)\,dx\Big]^{\frac{1}{m}} \qquad (m \geq m_1). \tag{19}$$

(18), (19) prove our statement, for they lead to the inequalities

$$\overline{\lim_{m\to\infty}}\, i_m^{\frac{1}{m}} \leq F, \qquad \underline{\lim_{m\to\infty}}\, i_m^{\frac{1}{m}} \geq F.$$

We shall not dwell here upon possible generalizations of theorem III.[9])

---

[9]) In theorems II and III the integrals involved can be written as Stieltjes integrals, for $\int_a^b p(x)\,f(x)\,dx = \int_a^b f(x)\,d\psi(x)$, where $\psi(x) = \int_a^x p(x)\,dx$. — As an

(Fortsetzung der Fußnote auf S. 164.)

§ 5.

Hereafter, the interval (a, b) *is assumed to be finite.* We proceed now to set forth an upper limit for $|f(x) - P_{nm}(x)|$, under the assumption: *$f(x)$ is continuous for $a \leqq x \leqq b$.* Introduce its polynomial of the best approximation on (a, b) $\Pi_n(x)$, of degree $\leqq n$, and the best approximation $E_n(f)$ (see Introduction):

$$(20) \qquad E_n(f) = \max |f(x) - \Pi_n(x)| \leqq \max |f(x) - G_n(x)| \quad \text{in } (a, b).$$

We have then, by the definition of $P_{nm}(x)$ and by (4),

$$(21) \qquad \begin{aligned} &\int_a^b p(x)\,|f(x) - P_{nm}(x)|^m\,dx \\ &= I_{nm} \leqq \int_a^b p(x)\,|f(x) - \Pi_n(x)|^m\,dx \leqq E_n^m(f) \int_a^b p(x)\,dx, \end{aligned}$$

$$(22) \qquad \int_a^b p(x)\,|P_{nm}(x) - \Pi_n(x)|^m\,dx \leqq [2\,E_n(f)]^m \int_a^b p(x)\,dx.$$

Assume, first, m, not being an integer, is > 2, and denote by 2μ the greatest even integer contained in m. By (4), we get from (22),

$$(23) \qquad \int_a^b p(x)\,|P_{nm}(x) - \Pi_n(x)|^{2\mu}\,dx \leqq [2\,E_n(f)]^{2\mu} \int_a^b p(x)\,dx,$$

and this inequality evidently holds also for

$$(24) \qquad 2\mu \leqq m < 2\mu + 2.$$

Apply (9) to the polynomial $[P_{nm}(x) - \Pi_n(x)]^\mu$ of degree $\leqq \mu n$, taking $m = 2$ and replacing n by μn. We get,

$$(25) \qquad |P_{nm}(x) - \Pi_n(x)| \leqq 2E_n(f)\Big[K_{n\mu}(x) \int_a^b p(t)\,dt\Big]^{\frac{1}{2\mu}} \qquad (m \geqq 2;\ x \text{ arbitrary}),$$

$$(26) \qquad |f(x) - P_{nm}(x)| \leqq E_n(f)\Big\{1 + 2\Big[K_{n\mu}(x) \int_a^b p(t)\,dt\Big]^{\frac{1}{2\mu}}\Big\}.$$

Formulae (25), (26) *hold for any c-function $p(x)$, for any continuous function $f(x)$.* [For $1 \geqq m > 2$ we could use (11).] They yield

---

illustration to theorem II may serve: 1°. The integral $c_n = \int_a^b p(x)\,|x|^n\,dx$, (a, b) finite. Here $\lim\limits_{n \to \infty} c_n^{\frac{1}{n}} \left(= \lim \frac{c_{n+1}}{c_n}\right) = \max(|a|, |b|)$. Cf. O. Perron, Die Lehre von den Kettenbrüchen (1913), p. 3–520, p. 384–385; 2°. $\lim\limits_{n \to \infty} \Big|\int_0^\infty e^{-nx^\lambda} x^n\,dx\Big]^{\frac{1}{n}}$ $(\lambda > 0)$ $= \frac{1}{\lambda}\, e^{-\frac{1}{\lambda^2}}$

the order, with respect to n, of $|f(x)-P_{nm}(x)|$ for very general $p(x)$, where the order of $K_n(x) \equiv \sum_{i=0}^{n} \varphi_i^2(x)$ is known.

Assume, to illustrate, that $p(x)$ satisfies the following

Condition P. *There exists an interval* (c, d) $(a \leqq c < d \leqq b)$ *with the following properties*: α) *it contains a finite number of points* x_i *to which correspond sufficiently small intervals* $(x_i-\delta_i,\ x_i+\delta_i)$ (the interval being one-sided if $x_i = c, d$) *in which* $\frac{p(x)}{|x-x_i|^{k_i}} > A_i$ *with certain* $k_i > -1$, $A_i > 0$; β) $p(x) \geqq p_0 > 0$ *elsewhere in* (a, b).

Then, as it has been shown by the writer[10]),

$$K_n(x) < \tau n^{\sigma} \qquad (c \leqq x \leqq d), \tag{27}$$

σ (>0, determined by the k_i) and τ do not depend on x, nor on n.

Therefore, *with* $p(x)$ *satisfying the condition* (P),

$$|f(x)-P_{nm}(x)| < \tau_m E_n(f)\, n^{\frac{\sigma}{2\mu}} \qquad (c \leqq x \leqq d), \tag{28}$$

where τ_m does not depend on n, nor on x, nor on $f(x)$, and this can be rewritten, as

$$|f(x)-P_{nm}(x)| < \tau_{nm} E_n(f)\, n^{\frac{\sigma}{\mu}} \qquad (c \leqq x \leqq d), \tag{29}$$

in case

$$n, m \to \infty, \quad \frac{\log n}{m^2} \text{ remains finite.} \tag{30}$$

With the additional assumption

$$p(x) \geqq p_0 > 0 \quad \text{for} \quad (a \leqq) c \leqq x \leqq d (\leqq b) \tag{31}$$

we can derive an upper limit for $|f(x)-P_{nm}(x)|$ by an entirely different method.

It makes no use of Tchebycheff polynomials, being based upon the well known Markoff-Bernstein theorem which we state as follows:

$|G_n(x)| \leqq M$ *on* (α, β) *implies*: 1°. $|G_n'(x)| \leqq Cn^2 M$ $(\alpha \leqq x \leqq \beta)$, 2°. $|G_n'(x)| \leqq CnM$ $(\alpha+\varepsilon \leqq x \leqq \beta-\varepsilon)$, *where* C (of course, not the same in 1°, 2°) *depends on* α, β *and* ε *only*.

We write, with the notations used above,

$$[P_{nm}(x)-\Pi_n(x)]^{2\mu} = \frac{d}{dx}\int_c^x [P_{nm}(t)-\Pi_n(t)]^{2\mu} dt.$$

---

[10]) J. Shohat, On the development of continuous functions in series of Tchebycheff polynomials, Transactions of the American Mathematical Society 27 (1925) p. 537–550, p. 540–541.

The polynomial $Q(x) = \int\limits_{c+\varepsilon}^{x} [P_{nm}(t) - \Pi_n(t)]^{2\mu} dt$, of degree $\leqq 2n\mu + 1$, satisfies, for $c \leqq x \leqq d$, the inequality (see (23))

$$|Q(x)| \leqq \frac{1}{p_0} \int\limits_a^b p(t)\, [P_{nm}(t) - \Pi_n(t)]^{2\mu} dt \leqq \frac{[2E_n(f)]^{2\mu}}{p_0} \int\limits_a^b p(t)\, dt,$$

and Markoff-Bernstein theorem gives:

$$
\begin{aligned}
&|P_{nm}(x) - \Pi_n(x)| < \tau_{nm} n^q E_n(f),\\
(32)\qquad &|f(x) - P_{nm}(x)| \ \ < \tau_{nm} n^q E_n(f),\\
&q = 2 \ \text{ for } \ c \leqq x \leqq d; \quad q = 1 \ \text{ for } \ c + \varepsilon \leqq x \leqq d - \varepsilon.^{11)}
\end{aligned}
$$

This is a special case of (25), (26), for, under (31),

(33) $K_n(x) < \tau n^2 \ (c \leqq x \leqq d), \quad K_n(x) < \tau n \ (c + \varepsilon \leqq x \leqq d - \varepsilon).$[10]

In addition to (31) we subject now $p(x)$ to the condition

$$(34) \qquad \int\limits_x^{x+\delta} p(x)\, dx < K\delta \qquad (c \leqq x < x + \delta \leqq d),$$

where K does not depend on x, nor on δ.

Then we get a new expression for the upper limit of $|f(x) - P_{nm}(x)|$ involving the modulus of continuity $\omega(h)$ of $f(x)$ in (a, b), which proves to be especially useful in case: n is fixed, $m \to \infty$. We write:

$$f(x) - P_{nm}(x) = \frac{1}{h} \int\limits_x^{x+h} [f(x) - f(t)]\, dt + \frac{1}{h} \int\limits_x^{x+h} [f(t) - P_{nm}(t)]\, dt$$

$$+ \frac{1}{h} \int\limits_x^{x+h} [P_{nm}(t) - P_{nm}(x)]\, dt \equiv i_1 + i_2 + i_3 \quad (h > 0;\ c \leqq x < x + h \leqq d).$$

$$
\begin{aligned}
&|i_1| \leqq \omega(h),\\
&|i_2| \leqq \frac{1}{hp_0} \Big[\int\limits_a^b p(t)\, |f(t) - P_{nm}(t)|^m dt\Big]^{\frac{1}{m}} \Big[\int\limits_x^{x+h} p(t)\, dt\Big]^{\frac{m-1}{m}} \leqq \frac{K^{\frac{m-1}{m}} h^{-\frac{1}{m}} E_n(f)}{p_0},\\
&|i_3| \leqq C n^q h M_{nm}
\end{aligned}
$$

(by Markoff-Bernstein theorem; $M_{nm} = \max |P_{nm}(x)|$ in (c, d)).

---

11) The elegant method of D. Jackson, also based essentially on Markoff-Bernstein theorem (*loc. cit.* (1-b), p. 162–164), seems to be inapplicable here directly, unless an additional assumption be made concerning the lower limit of $\Big|\int\limits_x^{x\pm\delta} p(x)\, dx\Big|$, $\delta = \dfrac{b-a}{8n^2}$, x, $x \pm \delta$ *varying throughout the whole interval* (a, b).

$$(35)\qquad R_{nm}(x) \equiv |f(x) - P_{nm}(x)| \leqq \omega(h) + \frac{K^{\frac{m-1}{m}} h^{-\frac{1}{m}} E_n(f)}{p_0} + Cn^q M_{nm} h$$

$$(q \text{ given in } (32);\ c \leqq x \leqq d).$$

$$|P_{nm}(x)| \leqq R_{nm}(x) + |f(x)|.$$

$$(36)\qquad \begin{aligned} M_{nm} &= |P_{nm}(\xi)| \leqq R_{nm}(\xi) + F \leqq R_{nm} + F \qquad (\xi \text{ in } (c,d)),\\ R_{nm} &= \max|R_{nm}(x)|, \quad F = \max|f(x)| \quad \text{in } (c,d). \end{aligned}$$

(35), (36) lead to

$$(37)\qquad |f(x) - P_{nm}(x)| \leqq \left\{\omega(h) + \frac{K^{\frac{m-1}{m}} h^{-\frac{1}{m}} E_n(f)}{p_0} + Cn^q h F\right\} : (1 - Cn^q h)$$

$$(Cn^q h < 1),$$

$$(38)\qquad |f(x) - P_{nm}(x)| \leqq E_n(f)\left\{2 + \frac{K^{\frac{m-1}{m}} h^{-\frac{1}{m}}}{p_0} + Cn^q h\right\} : (1 - Cn^q h),$$

$$(39)\qquad |f(x) - P_{nm}(x)| < 2E_n(f)\left\{3 + \frac{K^{\frac{m-1}{m}} h^{-\frac{1}{m}}}{p_0}\right\}.$$

(We derive (38), (39), applying (37) to $\varphi(x) \equiv f(x) - P_{nm}(x)$ and taking $Cn^q h = \frac{1}{2}$).

In (37), (38), (39) $q = 2$ *for* $c \leqq x \leqq d$, $q = 1$ *for* $c + \varepsilon \leqq x \leqq d + \varepsilon$, C *is given by Markoff-Bernstein theorem.*

The expressions given above for the upper limit of $|f(x) - P_{nm}(x)|$ enable us to treat I_{nm} and $P_{nm}(x)$ in all the cases given in the Introduction. Formula (26) also yields some interesting applications to the theory of Tchebycheff polynomials.

§ 6.

Case I. m is fixed, $n \to \infty$.

Theorem IV. *The sequence* $\left\{I_{nm}^{\frac{1}{m}}\right\}$ *is monotonically decreasing towards zero for any* $f(x)$ *of the class* $[L_p^m]$ $(m \geqq 1)$.

Proof. We have, first, by the definition of $P_{nm}(x)$,

$$(40)\qquad I_{nm} \geqq I_{n+1,m} \geqq I_{n+2,m} \geqq \cdots.$$

On the other hand, $f(x)$ being of the class $[L_p^m]$ $(m \geqq 1)$, a *continuous function* $\varphi(x)$ can be determined such that $\int_a^b p(x)|f(x) - \varphi(x)|^m dx < \frac{\varepsilon}{2^m}$.[12])

[12]) The proof is essentially the same as that given by Hobson (The Theory of functions of a real variable, 2-d ed., 2, p. 250) for the case $p(x) = 1$.

By Weierstraß' theorem, a polynomial $Q(x)$, of sufficiently high degree n can be assigned so that $\int_a^b p(x)|\varphi(x)-Q(x)|^m dx < \frac{\varepsilon}{2^m}$. Then, using (5) and the definition of $P_{nm}(x)$, we get,

$$\int_a^b p(x)|f(x)-P_{nm}(x)|^m dx \leqq \int_a^b p(x)|f(x)-Q(x)|^m dx$$

$$\leqq 2^{m-1}\left\{\int_a^b p(x)|f(x)-\varphi(x)|^m dx + \int_a^b p(x)|\varphi(x)-Q(x)|^m dx\right\} < \varepsilon,$$

which, combined with (40), proves our statement.

Assuming $m>1$, and writing $P_n(x), I_n$ in place of $P_{nm}(x), I_{nm}$, we associate with any $f(x)$ of the class $[L_p^m]$ the sequence of polynomials, uniquely determined,

$$(41) \qquad P_1(x),\ P_2(x),\ \ldots,\ P_n(x),\ \ldots$$

or, which is the same, the infinite series

$$(42) \qquad P_1(x)+[P_2(x)-P_1(x)]+\ldots+[P_{n+1}(x)-P_n(x)]+\ldots.$$

The sequence (41) may be spoken of as approximating $f(x)$ in the sense of the "least m-th powers" (D. Jackson), also — slightly generalizing a notion due to F. Riesz —, as "converging strongly to $f(x)$ with exponent m" [13]). We have then,

$$(43) \qquad \lim_{n\to\infty}\int_a^b p(x)|P_n(x)|^m dx = \int_a^b p(x)|f(x)|^m dx,$$

$$(44) \qquad \lim_{n\to\infty}\int_\alpha^t p(x)P_n(x)g(x)\,dx = \int_\alpha^t p(x)f(x)g(x)\,dx$$

$$\left(a\leqq\alpha,\ t\leqq b;\ g(x) \text{ of the class } \left[L_p^{\frac{m}{m-1}}\right]\right).$$

In other words, *for any* $f(x)$ *of the class* $[L_p^m]$ $(m>1)$ *the infinite series* (42), *multiplied on both sides by* $p(x)g(x)$, *where* $g(x)$ *is an arbitrary function of the class* $\left[L_p^{\frac{m}{m-1}}\right]$, *can be integrated term by term between any two limits* $(a\leqq)\alpha$, $t(\leqq b)$. *The convergence is uniform, if* t *be variable.*

The proof of (43), (44) is essentially the same as that given by Hobson[14]) for the case $p(x)=1$.

---

[13]) *Loc. cit.* [8]), p. 464. Here we introduce $p(x)$ as a factor.

[14]) *Loc. cit.* [13]), p. 251.

With regard to the convergence of the series (42), we derive from (26)

Theorem V. *The series* (42), *or, which is the same, the sequence* $\{P_{nm}(x)\}$, *with* m (>1) *fixed and* $n \to \infty$, *converges to* $f(x)$, *assumed to be continuous in* (a, b), *uniformly over any sub-interval* (c, d) $(a \leqq c < d \leqq b)$ *where* $K_n(x) = \sum_{i=0}^{n} \varphi_i^2(x) = O(n^\sigma)$ (this certainly takes place if the condition (P) is satisfied), *provided,* $\lim_{n \to \infty} E_n(f) n^{\frac{\sigma}{2\mu}} = 0$ (σ does not depend on n, nor on x).

From (31) we derive as sufficient condition for the uniform convergence of (41):

$$\lim_{n \to \infty} E_n(f) n^{\frac{q}{m}} = 0 \tag{45} \qquad \text{(see (32))}$$

quite similar to that given by D. Jackson in the case $p(x) = 1$.[15])

§ 7.

Case II. n is fixed, $m \to \infty$.

Consider the class of all $f(x)$ *measurable and bounded on* (a, b). Denote by $\overline{F}$, F respectively the upper bound and the "measurable upper bound" of $|f(x)|$ in (a, b).

The notion of the "best approximation" of functions by means of polynomials of degree $\leqq n$, established by Tchebycheff for functions *continuous* over a finite interval, can be extended to the more general class *of measurable bounded functions* $f(x)$ (under our consideration) *in two ways*, as follows. Using for the best approximation and the corresponding polynomial resp. the notations $E_n^\alpha(f)$, $E_n^\beta(f)$, $\Pi_n^\alpha(x)$, $\Pi_n^\beta(x)$, we define:

$E_n^\alpha(f) =$ *upper bound of* $|f(x) - \Pi_n^\alpha(x)|$, $E_n^\beta(f) =$ *measurable upper bound of* $|f(x) - \Pi_n^\beta(x)|$ *on* (a, b), *are each the smallest possible among all such expessions formed with an arbitrary polynomial of degree* $\leqq n$.

The existence of one at least $\Pi_n^\alpha(x)$ has been proved by Kirchberger[16]). In a similar manner we prove the existence of one at least $\Pi_n^\beta(x)$. The latter, which seems to be new, is more important, as is shown by theorems VI, VII below.

If $f(x)$ be continuous on (a, b) then

$$\overline{F} = F, \quad \Pi_n^\alpha(x) = \Pi_n^\beta(x) = \Pi_n(x), \quad E_n^\alpha(f) = E_n^\beta(f) = E_n(f).$$

---

[15]) *Loc. cit.* [1]) b), p. 165–166.

[16]) P. Kirchberger, Über Tchebycheffsche Annäherungsmethoden, Math. Annalen 57 (1903), p. 509–540; p. 511–512.

In general,

$$\overline{F} + F, \quad \Pi_n^\alpha(x) + \Pi_n^\beta(x), \quad E_n^\alpha(f) + E_n^\beta(f),$$

and we have, denoting respectively by F^α, $\overline{F}^\beta$ the measurable upper bound of $|f(x) - \Pi_n^\alpha(x)|$ and the upper bound of $|f(x) - \Pi_n^\beta(x)|$, for $a \leqq x \leqq b$,

$$(46) \qquad \begin{aligned} &E_n^\beta(f) \leqq F^\alpha \leqq E_n^\alpha(f) \leqq \overline{F}^\beta, \\ &E_n^\alpha(f) \geqq E_{n+1}^\alpha(f) \geqq \dots; \quad E_n^\beta(f) \geqq E_{n+1}^\beta(f) \geqq \dots . \end{aligned}$$

There exist, therefore,

$$\lim_{n \to \infty} E_n^\alpha(f) \geqq 0, \quad \lim_{n \to \infty} E_n^\beta(f) = E^\beta(f) \geqq 0.$$

Theorem VI. 1°. $\lim\limits_{n \to \infty} E_n^\alpha(f) = 0$, *if* $f(x)$ *has one at least discontinuity in* (a, b). 2°. $\lim\limits_{n \to \infty} E_n^\beta(f) = 0$ *for all* $f(x)$ *measurable and bounded in* (a, b).

Proof. 1°. The assumption $\lim\limits_{n \to \infty} E_n^\alpha(f) = 0$ implies:

$$|f(x) - \Pi_n^\alpha(x)| \leqq E_n^\alpha(f) < \varepsilon \qquad (a \leqq x \leqq b;\ n \geqq n_0);$$

in other words, the infinite series of polynomials

$$\Pi_1^\alpha(x) + [\Pi_2^\alpha(x) - \Pi_1^\alpha(x)] + \dots$$

converges to $f(x)$ uniformly for $a \leqq x \leqq b$, which necessitates the continuity of $f(x)$ throughout the whole interval (a, b), contrary to our hypothesis.

2°. Assume $E^\beta(f) > 0$. Denote by Γ_n the measurable upper bound on (a, b) of $|f(x) - G_n(x)|$. Then, for any n,

$$\Gamma_n \geqq E_n^\beta(f) \geqq E^\beta(f) > 0.$$

This contradicts a theorem of Hobson[17]) which states the existence of a sequence of polynomials $Q(x)$ converging, as n, the degree of $Q(x)$, $\to \infty$, to $f(x)$ almost everywhere in (a, b).

§ 8.

Let us associate with the polynomial $G_n(x) = \sum\limits_{i=0}^{n} g_i x^i$ the point $G_n = (g_0, g_1, \dots, g_n) = \left(\overset{n}{\underset{0}{g_i}}\right)$ in the n-dimensional space. *The relation* $A_n \to B_n$ *means, then, that the coefficients of* $A_n(x) = \sum\limits_{i=0}^{n} a_i x^i$ *converge respectively to those of* $B_n(x) = \sum\limits_{i=0}^{n} b_i x^i$, *and thus* $A_n(x) \to B_n(x)$ *uniformly over any finite interval.*

---

[17]) *Loc. cit.* [15]), p. 256.

Theorem VII. *Let the c-function $p(x)$ satisfy the requirements of theorem* II, *and let $f(x)$ be measurable and bounded on (a, b). Then:* 1°. *All limit-points of the set $\{P_{nm}(x)\}$ are in a finite portion of space and coincide with the set of the polynomials of the best approximation $\Pi_n^\beta(x)$ (n is fixed, $m \to \infty$).* 2°. $\lim\limits_{m\to\infty} I_{nm}^{\frac{1}{m}} = E_n^\beta(f)$. 3°. *If $f(x)$ be continuous on (a, b), then* $\lim\limits_{m\to\infty} P_{nm}(x) = \Pi_n(x)$, $\lim\limits_{m\to\infty} I_{nm}^{\frac{1}{m}} = E_n(f)$. 4°. *If m increases monotonically, so does* $\left[\frac{I_{nm}}{\int_a^b p(x)\,dx}\right]^{\frac{1}{m}}$, *and* $\lim\limits_{m\to\infty} I_{nm}^{\frac{1}{m}}$ *exists for any c-function $p(x)$.*

Proof. 1°. 2°. F denoting, as above, the measurable upper bound of $|f(x)|$ in (a, b),

$$(47)\qquad \int_a^b p(x)\,|f(x) - P_{nm}(x)|^m\,dx < (F+\varepsilon)^m \int_a^b p(x)\,dx \qquad \text{(see (13))}.$$

It follows (making use of the Corollary to theorem I) that *the coefficients of all $P_{nm}(x)$ are bounded, as functions of m*, which proves the first part of 1°. To prove the second part, we employ an argument somewhat similar to that of G. Pólya[18]). Let $Q_n \equiv \left(\begin{smallmatrix} n \\ q_i \\ 0 \end{smallmatrix}\right)$ denote *any one* of the limit-points in question. Denote, further, by Q the measurable upper bound of $|f(x) - Q_n(x)|$ in (a, b), and by E_ε the set of points x in (a, b) where

$$(48)\qquad |f(x) - Q_n(x)| \leqq Q - \frac{\varepsilon}{2}.$$

There exists in the set $\{m\}$ a sequence m_i ($i = 1, 2, \ldots$; $\lim\limits_{i\to\infty} m_i = \infty$) such that $\lim\limits_{i\to\infty} P_{nm_i}(x) = Q_n(x)$ uniformly in (a, b), and

$$|\{f(x) - P_{nm_i}(x)\} - \{f(x) - Q_n(x)\}| < \frac{\varepsilon}{2} \qquad (a \leqq x \leqq b;\ i \geqq i_0),$$

$$|f(x) - P_{nm_i}(x)| \geqq Q - \varepsilon \qquad (x \text{ in } E_\varepsilon;\ i \geqq i_0;\ \text{see (48)}),$$

$$(49)\qquad I_{nm_i}^{\frac{1}{m_i}} \geqq (Q-\varepsilon)\left[\int_{E_\varepsilon} p(x)\,dx\right]^{\frac{1}{m_i}} \qquad (i \geqq i_0).$$

By the definition of $P_{nm}(x)$, we write,

$$(50)\qquad I_{nm'}^{\frac{1}{m'}} \leqq I_{nm''}^{\frac{1}{m''}} \left[\int_a^b p(x)\,dx\right]^{\frac{m''-m'}{m'm''}} \qquad (m'' > m') \qquad \text{(see (4))}$$

$$(51)\qquad I_{nm}^{\frac{1}{m}} \leqq \left[\int_a^b p(x)\,|f(x) - \Pi_n^\beta(x)|^m\,dx\right]^{\frac{1}{m}} < (E_n^\beta(f)+\varepsilon)\left[\int_a^b p(x)\,dx\right]^{\frac{1}{m}}.$$

---

[18]) *Loc. cit.* [1]) a).

We assume, further,

(52) $$\int_a^b p(x)\,dx = 1,$$

replacing, if necessary, $p(x)$ by $cp(x)$, with $c > 0$ properly chosen. Then,

(53) $$I_{nm}^{\frac{1}{m}} \geqq I_{nm_i}^{\frac{1}{m_i}} \geqq (Q-\varepsilon)\left[\int_{E_\varepsilon} p(x)\,dx\right]^{\frac{1}{m_i}} \quad (m > m_i;\ i \geqq i_0) \text{ (see (49), (50))}.$$

(51), (53), where ε does not depend on m and $\lim\limits_{i\to\infty} m_i = \infty$, lead successively, making use of the definition of $\Pi_n^\beta(x)$, $E_n^\beta(f)$, to

(54) $$\overline{\lim_{m\to\infty}}\, I_{nm}^{\frac{1}{m}} \leqq E_n^\beta(f), \quad \underline{\lim_{m\to\infty}}\, I_{nm}^{\frac{1}{m}} \geqq Q,$$

(55) $$Q = E_n^\beta(f) = \overline{\lim_{m\to\infty}}\, I_{nm}^{\frac{1}{m}} = \underline{\lim_{m\to\infty}}\, I_{nm}^{\frac{1}{m}};\quad \lim_{i\to\infty} P_{nm_i}(x) = Q_n(x) \equiv \Pi_n^\beta(x).$$

(55) *holds even if we reject* (52), for replacing $p(x)$ by $cp(x)$ leads to cI_{nm} with $\lim\limits_{m\to\infty} c^{\frac{1}{m}} = 1$, while $P_{nm}(x)$ *remains unchanged.*

3° follows from 1° and 2°, $\Pi_n^\beta(x)$ here being *unique* and $\equiv \Pi_n(x)$, with $E_n^\beta(f) = E_n(f)$.

4° is proved by (50).

Theorem VII supplements and generalizes the results previously given by G. Pólya and the writer. The condition $\int_E p(x)\,dx > 0$ $(mE > 0)$ is indispensable if we want theorem VII, 3° to be valid for *all* continuous functions.

With such functions we can go further, if we assume

(56) $$p(x) = \text{const. for } (a \leqq) c \leqq x \leqq d (\leqq b).$$

Applying (37) with $K = p_0$, we get, writing, in general, $E_n(f;\, a, b)$ and taking

(57) $$h = (1 + m^{-\theta})^{-m} \quad (0 < \theta < 1), \quad h \to 0, \quad \text{as } m \to \infty:$$

(58) $$\begin{gathered} E_n(f; c, d) \leqq |f(x) - P_{nm}(x)| \leqq E_n(f; a, b) + \eta(m) \quad (c \leqq x \leqq d) \\ \eta(m) = O(m^{-\theta}) + \omega(e^{-m^{1-\theta}}), \quad \eta(m) \to 0, \quad \text{as } m \to \infty. \end{gathered}$$

Formula (58), in case $c = a$, $d = b$, is another proof of the result of G. Pólya: $\lim\limits_{m\to\infty} P_{nm}(x) = \Pi_n(x)$. It gives, moreover, some indication about the rapidity of this limiting process.

§ 9.

Case III. $n, m \to \infty$.

Writing m_n, I_n, $P_n(x)$, we state

Theorem VIII. 1°. *$I_n^{\frac{1}{m_n}}$ remains finite, as $n \to \infty$, for all $f(x)$ with finite measurable bounds in (a, b). 2°. α) $\lim\limits_{n \to \infty} I_n = 0$ for all $f(x)$ measurable and bounded in (a, b). β) For all such $f(x)$ the infinite series $P_\nu(x) + [P_{\nu+1}(x) - P_\nu(x)] + \ldots$, where ν is sufficiently large so that $m_\nu > 1$, multiplied by $p(x) g(x)$ and integrated term by term between the limits a, t $(a \leqq a,\ t \leqq b)$ converges uniformly to $\int\limits_a^t p(x) f(x) g(x)\, dx$, provided, $\int\limits_a^b p(x) |g(x)|^{1+r} dx$ exists with a certain $r > 0$.*

Proof. 1°. F denoting the measurable upper bound of $|f(x)|$ in (a, b), we write, by the definition of $P_n(x)$,

$$I_n < (F + \varepsilon)^{m_n} \int\limits_a^b p(x)\, dx,$$

and this proves our statement, since, as $n \to \infty$, $m_n \to \infty$, $\left[\int\limits_a^b p(x)\, dx\right]^{\frac{1}{m_n}} \to 1$.

2°. α). With $\Pi_n^\beta(x)$ and $E_n^\beta(f)$ introduced above, we write,

$$I_n^{\frac{1}{m_n}} \leqq \left[\int\limits_a^b p(x) |f(x) - \Pi_n^\beta(x)|^{m_n} dx\right]^{\frac{1}{m_n}} < (E_n^\beta(f) + \varepsilon) \left[\int\limits_a^b p(x)\, dx\right]^{\frac{1}{m_n}},$$

with $\lim\limits_{n \to \infty} E_n^\beta(f) = 0$.

β) $\left|\int\limits_a^b p(x) g(x) (f(x) - P_n(x))\, dx\right|$

$$\leqq \left[\int\limits_a^b p(x) |f(x) - P_n(x)|^{m_n} dx\right]^{\frac{1}{m_n}} \cdot \left[\int\limits_a^b p(x) |g(x)|^{\frac{m_n}{m_n - 1}} dx\right]^{\frac{m_n - 1}{m_n}},$$

and the right-hand member $\to 0$, as $n \to \infty$, for so does the first factor (by 2°, α)), while the second factor exists as a finite number for all n sufficiently large so that $\frac{m_n}{m_n - 1} < 1 + r$.

The most interesting case is that of $f(x)$ *continuous in* (a, b).

Theorem IX. *Given a sequence of exponents m_n such that $\lim\limits_{n \to \infty} m_n = \infty$, $\frac{\log n}{m_n}$ remains finite, and a c-function $p(x)$ defined over a finite interval (a, b). Then, $f(x)$ being an arbitrarily given continuous on (a, b)*

function, the polynomial $P_n(x)$, *of degree* $\leqq n$, *minimizing the integral* $\int_a^b p(x)|f(x)-P_n(x)|^{m_n}dx$ *converges, as* $n\to\infty$, *to* $f(x)$ *uniformly over any sub-interval* (c, d) $(a\leqq c<d\leqq b)$, *where* $\sum_{i=0}^{n}\varphi_i^2(x)=O(n^\sigma)$ $(\sigma>0$ *independent on* $n, x)$. *Furthermore, in case* $c=a$, $d=b$, *the approximation of* $f(x)$ *by* $P_n(x)$ *on* (a, b) *is of the same order with repect to* n, *taken sufficiently large, as the best approximation* $E_n(f)$.

This follows at once from (29), since $\max|f(x)-P_n(x)|$ on (a, b) $\geqq E_n(f)$.

It suffices, therefore, to take $m_n=2n$, $p(x)=(x-a)^{\alpha-1}(b-x)^{\beta-1}$ $(\alpha, \beta>0)$, or, simply, $p(x)=1$, and *we thus obtain a sequence of polynomials* $\{P_n(x)\}$, *following a simple formula, which converges to* $f(x)$, *subjected to the only condition of continuity, uniformly over the whole interval* (a, b), *the approximation being of the same order, with respect so* n, *as the best approximation.*

Corollary. *The approximation properties of* $P_n(x)$ *stated above hold for any sequence of exponents* $\{m_n\}$ *with* $\lim_{n\to\infty} m_n=\infty$, *provided,* $f(x)$ *satisfies in* (a, b) *Lipschitz's condition of an arbitrarily given order.*

In fact, such a condition of order, say, α implies:

$$
(59)\qquad \begin{aligned} &E_n(f)=O(n^{-\alpha}),\\ &E_n(f)\,n^{\frac{2}{m_n}}=O\left(n^{\frac{2}{m_n}-\alpha}\right)\to 0, \text{ as } n\to\infty. \end{aligned}
$$

The condition: $\frac{\log n}{m_n}$ remains finite for $n\to\infty$, is sufficient to assure the uniform convergence of $P_n(x)$ to $f(x)$ on (a, b). It would be interesting to find a condition relating to the mode of increase of the sequence $\{m_n\}$, which is *necessary* for such a convergence.

§ 10.
Application to Tchebycheff polynomials.

Consider our minimum problem $I_{nm}=\int_a^b p(x)|f(x)-P_{nm}(x)|^m dx$ $=\min\int_a^b p(x)|f(x)-G_n(x)|^m dx$ *for* $m=2$. Then, as it is well known,

$$
(60)\qquad P_{n,2}(x)=\sum_{i=0}^{n}A_i\varphi_i(x),\qquad A_i=\int_a^b p(x)f(x)\varphi_i(x)dx,
$$

and formula (44) gives,

$$
(61)\quad \int_a^t p(x)f(x)g(x)dx=\sum_{n=0}^{\infty}A_nB_n,\ B_n=\int_a^t p(x)g(x)\varphi_n(x)dx\ (a\leqq \alpha, t\leqq b),
$$

the convergence being uniform, with respect to t, for any $g(x)$ of the class $[L_p^2]$. In particular,

$$(62)\qquad \int_a^b p(x) f^2(x)\,dx = \sum_{n=0}^{\infty} A_n^2, \qquad A_n = \int_a^b p(x) f(x) \varphi_n(x)\,dx.$$

Formulae (61), (62) represent the so-called "closure-property" for Tchebycheff polynomials. *Thus, formula* (44) *extends this property to the minimizing polynomials* $P_{nm}(x)$ $(n = 1, 2, \ldots;\ m > 1)$.

Theorem V, combined with (33), (59), (60), leads to

Theorem X. *The infinite series*

$$\sum_{n=0}^{\infty} \varphi_n(x) \int_a^b p(x) f(x) \varphi_n(x)\,dx$$

converges to $f(x)$ *uniformly in any sub-interval* (c, d) $(a \leqq c,\ d \leqq b)$, *where* $p(x) \geqq p_0 > 0$, *provided,* $f(x)$ *has a continuous derivative in* (a, b). *A sufficient condition for this uniform convergence to hold inside* (c, d) *is:* $f(x)$ *satisfies Lipschitz's condition of order* $> \frac{1}{2}$ *in the interval* (a, b).

Proof. The following remark is sufficient: for $f(x)$ having a continuous derivative in (a, b) $E_n(f) = o\left(\frac{1}{n}\right)$.

§ 11.

The results given above hold, mutatis mutandis, if we replace polynomials (of degree n) by limited trigonometric sums (of order n).

The University of Michigan.

(Eingegangen am 15. 11. 1928.)

Berichtigung

zu der Arbeit von L. Vietoris „Zum höheren Zusammenhang der kompakten Räume“, Math. Annalen **101**, S. 219–225.

Der in der obigen Arbeit S. 223 unmittelbar nach (3) stehende Absatz hat durch ein Versehen meinerseits einen von freundlicher Mitwirkung eines Mitlesers der Korrekturen herrührenden Wortlaut bekommen, welcher mir den Anschein gibt, als wollte ich in der Prioritätsfrage hinsichtlich der Schaffung des Dimensionsbegriffes und der Entstehung der dimensionstheoretischen Hauptsätze ein Urteil abgeben. Da mir dies hier ferne liegt, setze ich anstatt dieses Absatzes hiermit den folgenden Wortlaut:

„Zum Beweis stützen wir uns auf den bekannten Satz:“

L. Vietoris.

Über einen Satz von Herrn Esclangon.

Von

Edmund Landau in Göttingen.

Einleitung.

Alle Zahlen in dieser Arbeit seien reell, $n \geqq 2$ ganz, y und alle p Funktionen von x; alle Ableitungen in Endpunkten von Intervallen einseitig nach innen gemeint.

Satz 1. *Für* $x \geqq 0$ *sei*

$$\begin{aligned} &y \text{ beschränkt,} \\ &y^{(n)} + p_1 y^{(n-1)} + \ldots + p_n y \text{ beschränkt,} \\ &p_m \quad (1 \leqq m \leqq n) \text{ beschränkt.} \end{aligned}$$

Dann ist für $x \geqq 0$

$$y^{(m)} \quad (0 < m \leqq n) \text{ beschränkt.}$$

Dieser Satz ist neu. Herr Esclangon[1]) bewies einen schwächeren Satz, indem er überdies ausdrücklich die Existenz und Beschränktheit von $p_m^{(l)}$ für $1 \leqq m < n$, $1 \leqq l \leqq n - m$ und während des Beweises (unnötigerweise) stillschweigend[2]) noch mehr voraussetzte.

Da der Esclangonsche Wortlaut — er ist richtig — in neuerer Zeit zu wichtigen Anwendungen[3]) in der Theorie der fastperiodischen Funktionen

---

[1]) *Sur les intégrales bornées d'une équation différentielle linéaire* [Comptes rendus hebdomadaires des séances de l'Académie des Sciences, Paris, **160** (1915), S. 475–478]; *Nouvelles Recherches sur les Fonctions quasi-périodiques* [Annales de l'Observatoire de Bordeaux **16** (1921), S. 51–177], S. 149–154.

[2]) Bereits der von Herrn Esclangon auf S. 476 bzw. S. 150 eingeführte Differentialausdruck mit dem q_n existiert sonst nicht; bekanntlich existiert nicht zu jeder beschränkten Funktion ein unbestimmtes oder bestimmtes (Riemannsches oder Lebesguesches) Integral.

[3]) Bohr und Neugebauer, *Über lineare Differentialgleichungen mit konstanten Koeffizienten und fastperiodischer rechter Seite* [Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse, Jahrgang 1926, S. 8–22], S. 10 und S. 14.

führte, will ich einen einfacheren Beweis für ihn, nämlich alsbald für den weitergehenden Satz 1 angeben. Allerdings liegt dieser Beweis dem Kenner der ersten Seiten einer berühmten Hardy-Littlewoodschen[4]), zwei Jahre vor Herrn Esclangons erstgenannter Note erschienenen Arbeit sehr nahe.

Fügt man noch die Voraussetzung der Stetigkeit von $y^{(n)}$ hinzu, so ist der entstehende Satz überhaupt keiner Publikation wert. Denn, wer Hardy-Littlewoods[5]) Theorem 3(a) kennt, muß sich sagen: Wird

$$\omega(x) = \operatorname{Max}(1, \operatorname*{Max}_{0 \leqq z \leqq x} |y^{(n)}(z)|)$$

gesetzt, so ist mit von x freien P für $x \geqq 0$, $0 \leqq m \leqq n$

$$|y^{(m)}| \leqq P_1 \omega^{\frac{m}{n}},$$

also

$$|y^{(n)}| \leqq P_2 + \sum_{m=1}^{n} |p_m|\,|y^{(n-m)}| \leqq P_3\,\omega^{\frac{n-1}{n}},$$

$$\omega \leqq \operatorname{Max}\left(1, P_3\,\omega^{\frac{n-1}{n}}\right) \leqq P_4\,\omega^{\frac{n-1}{n}},$$

$$\omega \leqq P_4^n,$$

$$|y^{(m)}| \leqq P_1 P_4^m \leqq P_5 \quad \text{für} \quad 0 \leqq m \leqq n.$$

Auch der volle Wortlaut des Satzes 1 ließe sich durch Hinzufügung weniger Zeilen zu Hardy-Littlewood beweisen. Doch will ich den Beweis ohne Zitate auf jene Abhandlung führen und werde statt dessen den in vierfacher Hinsicht weitergehenden Satz 2 und den in dreifacher Hinsicht weitergehenden Satz 3 beweisen.

Erstens wird die Voraussetzung der Beschränktheit der p_m und des Differentialausdrucks durch weniger ersetzt;

zweitens beziehen sich die Sätze auf ein endliches Intervall;

drittens wird in Satz 2 (nicht aber in Satz 3) vom Differentialausdruck nur eine einseitige Ungleichung vorausgesetzt (und entsprechend bei $y^{(n)}$ behauptet);

viertens hat die ganze Sache nichts mit **linearen** Differentialausdrücken zu tun; statt

$$p_1 y^{(n-1)} + \cdots + p_n y$$

(wo ja das Glied $p_n y$ ohne weiteres gestrichen werden kann) lege ich ein

---

[4]) *Contributions to the Arithmetic Theory of Series* [Proceedings of the London Mathematical Society (2) 11 (1913), S. 411–478].

[5]) Loc. cit., S. 422.

beliebiges Polynom

$$\mathfrak{P} = \mathfrak{P}(y', \ldots, y^{(n-1)}) = \Sigma\, p_{l_1, \ldots, l_{n-1}}\, y'^{\,l_1} \ldots y^{(n-1)\,l_{n-1}}$$

zugrunde, wo die l_m $(0 < m < n)$ die verschiedenen Systeme mit

$$(1) \qquad \sum_{m=1}^{n-1} m\, l_m < n, \qquad l_m \geqq 0 \text{ und ganz,}$$

durchlaufen; zur Abkürzung werde

$$N = N(l_1, \ldots, l_{n-1}) = \frac{1}{n} \sum_{m=1}^{n-1} m\, l_m$$

gesetzt.

Satz 2. *Es gibt zwei nur von n abhängige positive Zahlen $\varepsilon(n)$ und $P(n)$ mit folgender Eigenschaft:*

Für $|x| \leqq 1$ sei

$$|y| \leqq 1.$$

Es sei für jedes System (l_m) mit (1), wenn

$$(2) \qquad \mu = \operatorname{Max}(1, y^{(n)}(0))$$

gesetzt wird,

$$(3) \qquad |p_{l_1, \ldots, l_{n-1}}| < \varepsilon \mu^{1-N}$$

und

$$(4) \qquad y^{(n)} + \mathfrak{P} < \varepsilon \mu.$$

Dann ist

$$|y^{(m)}(0)| \leqq P \quad \textit{für} \quad 0 < m < n$$

und

$$y^{(n)}(0) \leqq P.$$

Satz 3. *Es gibt zwei nur von n abhängige positive Zahlen $\varepsilon(n)$ und $P(n)$ mit folgender Eigenschaft:*

Für $|x| \leqq 1$ sei

$$|y| \leqq 1.$$

Es sei für jedes System (l_m) mit (1), wenn

$$(5) \qquad \mu = \operatorname{Max}(1, |y^{(n)}(0)|)$$

gesetzt wird,

$$(3) \qquad |p_{l_1, \ldots, l_{n-1}}| < \varepsilon \mu^{1-N}$$

und

$$(6) \qquad |y^{(n)} + \mathfrak{P}| < \varepsilon \mu.$$

Dann ist

$$|y^{(m)}(0)| \leqq P \quad \textit{für} \quad 0 < m \leqq n.$$

Satz 3 ist natürlich ein Spezialfall des Satzes 2, obgleich eine Voraussetzung, wegen $y^{(n)}(0) \leqq |y^{(n)}(0)|$, geringer und eine Behauptung schärfer ist. Denn unter den Voraussetzungen des Satzes 3 ist nur Satz 2 im Falle $y^{(n)}(0) \geqq 0$ auf y, im Falle $y^{(n)}(0) \leqq 0$ auf $-y$ und $-\mathfrak{P}(-y', \ldots, -y^{(n-1)})$ statt $\mathfrak{P}$ anzuwenden.

Satz 3 enthält natürlich den Satz 1. Denn unter dessen Voraussetzungen ist bei passendem A für $x \geqq 0$

$$|y| < A,$$

$$|y^{(n)} + p_1 y^{(n-1)} + \ldots + p_{n-1} y'| < A, \tag{7}$$

$$|p_m| < A \quad (1 \leqq m \leqq n-1).$$

Wird

$$r = \mathrm{Min}\left(1, \frac{\varepsilon}{A}, \frac{\varepsilon}{2}\right)$$

gesetzt, so genügt

$$\frac{y(rx)}{A} = Y(x)$$

für $x \geqq 0$ den Bedingungen

$$|Y| \leqq 1,$$

$$\left|\frac{1}{r^n} Y^{(n)} + \frac{p_1}{r^{n-1}} Y^{(n-1)} + \ldots + \frac{p_{n-1}}{r} Y'\right| < 1,$$

$$|Y^{(n)} + p_1 r Y^{(n-1)} + \ldots + p_{n-1} r^{n-1} Y'| < r^n,$$

und hierin ist

$$|p_m r^m| < Ar \leqq \varepsilon \quad \text{für} \quad 0 < m < n,$$

$$r^n \leqq r < \varepsilon.$$

Satz 3 ist wegen $\mu \geqq 1$ bei jedem $\xi \geqq 1$ auf $Z(x) = Y(\xi + x)$ (statt y), $\mathfrak{P} = \sum_{m=1}^{n-1} p_m(\xi + x) r^m Z^{(n-m)}$ anwendbar und liefert

$$|Y^{(m)}(\xi)| \leqq P \quad \text{für} \quad 0 < m < n,$$

wo P von ξ frei ist.

$$y^{(m)}(x) = \frac{A}{r^m} Y^{(m)}\left(\frac{x}{r}\right) \qquad (0 < m < n)$$

ist also für $x \geqq r$ beschränkt und (wegen der Stetigkeit) selbstverständlich für $0 \leqq x \leqq r$. Nach (7) ist also auch $y^{(n)}$ für $x \geqq 0$ beschränkt.

Da Satz 3 wesentlich einfacher zu beweisen geht, will ich zunächst im ersten Teil diesen schwächeren Satz beweisen und erst im zweiten Teil den vollen Satz 2.

Erster Teil.

Satz 3.

§ 1.

Hilfssätze.

Hilfssatz 1.*) *Es sei* $a > 0$, $b > 0$. *Für* $|x| \leqq \sqrt{\frac{a}{b}}$ *sei*

$$|y| \leqq a, \quad |y''| \leqq b.$$

Dann ist dort

$$|y'| \leqq 2\sqrt{ab}.$$

Beweis. 1. Es sei $a = b = 1$. Für $|x| \leqq 1$ ist bei passenden ξ_1, ξ_2 mit $-1 \leqq \xi_1 \leqq x \leqq \xi_2 \leqq 1$

$$y(x) - y(-1) = (1+x)\,y'(x) - \frac{(1+x)^2}{2}\,y''(\xi_1),$$

$$y(1) - y(x) \quad = (1-x)\,y'(x) + \frac{(1-x)^2}{2}\,y''(\xi_2),$$

also

$$2y'(x) = y(1) - y(-1) + \frac{(1+x)^2}{2}\,y''(\xi_1) - \frac{(1-x)^2}{2}\,y''(\xi_2),$$

$$2|y'(x)| \leqq 1 + 1 + \frac{(1+x)^2}{2} + \frac{(1-x)^2}{2} = 3 + x^2 \leqq 4,$$

$$|y'(x)| \leqq 2.$$

2. Im allgemeinen Fall genügt

$$Y(x) = \frac{1}{a}\,y\left(x\sqrt{\frac{a}{b}}\right)$$

für $|x| \leqq 1$ den Bedingungen

$$|Y(x)| \leqq 1, \quad |Y''(x)| = \left|\frac{1}{b}\,y''\left(x\sqrt{\frac{a}{b}}\right)\right| \leqq 1.$$

Also ist für $|x| \leqq 1$ nach 1.:

$$\left|\frac{1}{\sqrt{ab}}\,y'\left(x\sqrt{\frac{a}{b}}\right)\right| = |Y'(x)| \leqq 2,$$

folglich für $|x| \leqq \sqrt{\frac{a}{b}}$

$$|y'(x)| \leqq 2\sqrt{ab}.$$

---

*) Nicht neu. Vgl. meine Arbeit *Einige Ungleichungen für zweimal differentiierbare Funktionen* [Proceedings of the London Mathematical Society (2) **13** (1914), S. 43–49], S. 44, Fußnote *.

Hilfssatz 2. *Es sei* $0 < a \leqq b$. *Für* $|x| \leqq 1$ *sei*

$$|y| \leqq a, \qquad |y''| \leqq b.$$

Dann ist dort

$$|y'| \leqq 2\sqrt{ab}.$$

Beweis. Hilfssatz 1 mit $y(\xi + x)$, $|\xi| \leqq 1 - \sqrt{\frac{a}{b}}$.

Hilfssatz 3. *Es sei für* $|x| \leqq \frac{1}{\sqrt[n]{\omega}}$

$$|y| \leqq 1, \qquad |y^{(n)}| \leqq \omega.$$

Dann ist dort für $0 < m < n$

$$|y^{(m)}| \leqq 2^{2^n} \omega^{\frac{m}{n}}.$$

Beweis. 1. Es sei $\omega = 1$. Man setze

$$\chi = \operatorname{Max}\left(2^{2^n}, \operatorname*{Max}_{0<m<n} \operatorname*{Max}_{|x| \leqq 1} |y^{(m)}|\right).$$

Durch Induktion wird sich auf $|x| \leqq 1$ für $0 \leqq m < n$

$$(8) \qquad |y^{(m)}| \leqq 4\chi^{1-\frac{1}{2^m}}$$

ergeben.

Für $m = 0$ ist

$$|y^{(m)}| = |y| \leqq 1 < 4 = 4\chi^{1-\frac{1}{2^m}}.$$

Aus (8) mit einem m auf $0 \leqq m < n-1$ folgt (8) mit $m+1$, wenn Hilfssatz 2 auf $y^{(m)}$ statt y, $a = 4\chi^{1-\frac{1}{2^m}}$, $b = \chi$ angewendet wird. In der Tat ist erstens

$$0 < a = 2^2 \chi^{1-\frac{1}{2^m}} \leqq \chi^{\frac{1}{2^{n-2}}} \chi^{1-\frac{1}{2^m}} \leqq \chi = b,$$

zweitens nach (8)

$$|y^{(m)}| \leqq a,$$

drittens für $m < n-2$ nach Definition von χ, für $m = n-2$ nach Voraussetzung (man beachte $1 \leqq \chi$)

$$|y^{(m+2)}| \leqq \chi = b.$$

Hilfssatz 2 ergibt also

$$|y^{(m+1)}| \leqq 2\sqrt{4\chi^{1-\frac{1}{2^m}}\chi} = 4\chi^{1-\frac{1}{2^{m+1}}};$$

d. h. (8) gilt mit $m+1$ statt m.

(8) gilt also für $0 \leqq m < n$.

Folglich ist

$$\underset{0<m<n}{\mathrm{Max}}\ \underset{|x|\leqq 1}{\mathrm{Max}}\ |y^{(m)}| \leqq 4\chi^{1-\frac{1}{2^{n-1}}}.$$

Andererseits ist nach Definition von χ

$$2^{2^n} = 4\,(2^{2^n})^{1-\frac{1}{2^{n-1}}} \leqq 4\chi^{1-\frac{1}{2^{n-1}}}.$$

Nach Definition von χ ist also

$$\chi \leqq 4\chi^{1-\frac{1}{2^{n-1}}},$$

$$\chi \leqq 4^{2^{n-1}} = 2^{2^n};$$

nach Definition von χ ist also

$$|y^{(m)}| \leqq 2^{2^n} \quad \text{für} \quad 0 < m < n.$$

2. Ist ω beliebig, so werde 1. auf

$$y\left(\frac{x}{\sqrt[n]{\omega}}\right) = Y(x)$$

angewendet. Auf $|x| \leqq 1$ ist

$$|Y(x)| \leqq 1, \quad |Y^{(n)}(x)| = \frac{1}{\omega}\left|y^{(n)}\left(\frac{x}{\sqrt[n]{\omega}}\right)\right| \leqq 1,$$

also nach 1.:

$$|Y^{(m)}(x)| \leqq 2^{2^n} \quad \text{für} \quad 0 < m < n,$$

also für $|x| \leqq \frac{1}{\sqrt[n]{\omega}}$, $0 < m < n$,

$$|y^{(m)}(x)| = \omega^{\frac{m}{n}}\,|Y^{(m)}(x\sqrt[n]{\omega})| \leqq 2^{2^n}\omega^{\frac{m}{n}}.$$

§ 2.

Beweis von Satz 3.

Es sei $k = k(n)$ die Lösungszahl von (1),

$$P = P(n) = 2^{2^n},$$

$$\varepsilon = \varepsilon(n) = \frac{1}{4kP^n} \quad \left(< \frac{1}{4}\right).$$

Ich behaupte, daß diese Zahlen P und ε das in Satz 3 Behauptete leisten.

$y^{(n)}$ ist nach (3) und (6) für $|x| \leqq 1$ beschränkt.

Ich setze

$$\omega = \operatorname{Max}(1, \underset{|x| \leqq 1}{\text{obere Grenze von}} |y^{(n)}|).$$

Nach Hilfssatz 3 $\left(\text{man beachte } \frac{1}{\sqrt[n]{\omega}} \leqq 1\right)$ mit $y(\xi + x)$, $|\xi| \leqq 1 - \frac{1}{\sqrt[n]{\omega}}$, ist, da für $|x| \leqq 1$

$$|y| \leqq 1, \quad |y^{(n)}| \leqq \omega$$

ist, für $|x| \leqq 1$, $0 < m < n$

$$(9) \qquad |y^{(m)}| \leqq 2^{2^n} \omega^{\frac{m}{n}} = P\omega^{\frac{m}{n}}.$$

Jedes der k Glieder von $\mathfrak{P}$ ist nach (3) (da $\mu \leqq \omega$ wegen (5)) absolut

$$< \varepsilon \omega^{1-N} P^{l_1 + \ldots + l_{n-1}} \omega^N \leqq \varepsilon P^n \omega = \frac{\omega}{4k};$$

nach (6) ist also für $|x| \leqq 1$

$$|y^{(n)}| < k \frac{\omega}{4k} + \frac{\omega}{4} = \frac{\omega}{2},$$

$$\underset{|x| \leqq 1}{\text{obere Grenze von}} |y^{(n)}| \leqq \frac{\omega}{2}.$$

Nach Definition von ω ist also

$$(10) \qquad \omega = 1,$$

$$|y^{(n)}(0)| \leqq \frac{1}{2} < P$$

und nach (9) und (10)

$$|y^{(m)}(0)| \leqq P \quad \text{für} \quad 0 < m < n.$$

Zweiter Teil.

Satz 2.

§ 1.

Hilfssätze.

Hilfssatz 4. *Für* $0 \leqq x \leqq \frac{1}{2}$ *sei*

$$|y| \leqq 1, \quad y^{(n)} \leqq 1.$$

Dann ist

$$y^{(n-1)}(0) \geqq -P_1(n).$$

Beweis. Mit $h = \frac{1}{2(n-1)}$ ist

$$\sum_{\nu=0}^{n-1} (-1)^{n-1-\nu} \binom{n-1}{\nu} y(\nu h) = \int_0^h dx_1 \int_{x_1}^{x_1+h} dx_2 \ldots \int_{x_{n-2}}^{x_{n-2}+h} y^{(n-1)}(x_{n-1})\, dx_{n-1}.$$

Die linke Seite ist

$$\geqq -\sum_{\nu=0}^{n-1} \binom{n-1}{\nu} = -2^{n-1}.$$

Im Integral rechts ist nach dem Mittelwertsatz

$$y^{(n-1)}(x_{n-1}) \leqq y^{(n-1)}(0) + x_{n-1} \leqq y^{(n-1)}(0) + \frac{1}{2}.$$

Daher ist

$$-2^{n-1} \leqq h^{n-1}\left(y^{(n-1)}(0) + \frac{1}{2}\right),$$

$$y^{(n-1)}(0) \geqq -P_1(n).$$

Hilfssatz 5. *Es sei* $\tau \geqq \frac{1}{\sqrt[n]{\omega}}$. *Für* $|x| \leqq \tau$ *sei*

$$|y| \leqq 1, \quad y^{(n)} \leqq \omega.$$

Dann ist für $|x| \leqq \tau - \frac{1}{\sqrt[n]{\omega}}$, $0 < m < n$,

$$|y^{(m)}| \leqq P_2(n)\, \omega^{\frac{m}{n}}.$$

Beweis. Ohne Beschränkung der Allgemeinheit sei $\omega = 1$ $\Big($sonst betrachte man $y\left(\frac{x}{\sqrt[n]{\omega}}\right)\Big)$.

Nach Hilfssatz 4, angewendet auf $y(x+\xi)$ mit $-\tau \leqq \xi \leqq \tau - \frac{1}{2}$ und auf $(-1)^n y(-x+\xi)$ mit $-\tau + \frac{1}{2} \leqq \xi \leqq \tau$, ist

$$y^{(n-1)}(\xi) \geqq -P_1(n) \quad \text{für} \quad -\tau \leqq \xi \leqq \tau - \frac{1}{2}$$

und

$$-y^{(n-1)}(\xi) \geqq -P_1(n) \quad \text{für} \quad -\tau + \frac{1}{2} \leqq \xi \leqq \tau.$$

Für $|x| \leqq \tau - \frac{1}{2}$ ist also

$$|y^{(n-1)}(x)| \leqq P_1(n). \tag{11}$$

Im Fall $n = 2$ sind wir wegen $\tau - 1 < \tau - \frac{1}{2}$ am Ziel.

Im Fall $n > 2$ werde der Hilfssatz für $n-1$ als wahr angenommen. Es ist nach (11), wo $P_1(n) \geqq 2^{n-1}$ gewählt werden darf $\left(\text{dann ist } \tau - 1 \leqq \tau - \frac{1}{2} - \frac{1}{\sqrt[n-1]{P_1(n)}}\right)$, mit $n-1$ statt n, $\tau - \frac{1}{2}$ statt τ, $P_1(n)$ statt ω, für $|x| \leqq \tau - 1$, $0 < m \leqq n-2$,

$$(12) \qquad |y^{(m)}| \leqq P_2(n-1)\, P_1^{\frac{m}{n-1}}(n) \leqq P_2(n-1)\, P_1(n).$$

Für $|x| \leqq \tau - 1$, $0 < m < n$ ist schließlich nach (11) und (12)

$$|y^{(m)}| \leqq \operatorname{Max}(P_1(n), P_2(n-1)\, P_1(n)) = P_2(n).$$

§ 2.

Beweis von Satz 2.

Das $P_2(n)$ des Hilfssatzes 5 darf ich $\geqq 1$ gewählt denken. Alsdann setze ich, wenn k die Lösungszahl von (1) ist,

$$\varepsilon = \varepsilon(n) = \frac{1}{2^{n+1}\, k\, P_2^n(n)} \quad \left(\leqq \frac{1}{2^{n+1}}\right)$$

und

$$P(n) = P_2(n)\, 4^n.$$

Ich behaupte, daß diese Zahlen ε und P das in Satz 2 Behauptete leisten.

$y^{(n)}$ ist nach den Voraussetzungen (3) und (4) für $|x| \leqq 1$ nach oben beschränkt.

Ich setze für $0 \leqq t \leqq 1$

$$\omega(t) = \operatorname{Max}(4^n, \underset{|x| \leqq t}{\text{obere Grenze von}}\ y^{(n)}(x)),$$

$$\Omega(t) = \sqrt[n]{\omega(t)}.$$

Dann fallen $\omega(t)$ und $\Omega(t)$ mit wachsendem t nicht, und es ist

$$\Omega(t) \geqq 4$$

und nach (2)

$$(13) \qquad \omega(t) \geqq \omega(0) \geqq \mu.$$

Es sei $0 \leqq t \leqq \frac{3}{4}$, also

$$\tau = \tau(t) = t + \frac{1}{\Omega(t)} \leqq t + \frac{1}{4} \leqq 1.$$

Für $|x| \leqq \tau$ ist

$$|y| \leqq 1,$$

$$y^{(n)} \leqq \omega(\tau).$$

Nach Hilfssatz 5 $\Bigl($man beachte

$$\tau - \frac{1}{\Omega(\tau)} = t + \frac{1}{\Omega(t)} - \frac{1}{\Omega(\tau)} \geqq t\Bigr)$$

ist für $|x| \leqq t$, $0 < m < n$,

(14)
$$|y^{(m)}| \leqq P_2(n)\,\omega^{\frac{m}{n}}(\tau) = P_2(n)\,\Omega^m(\tau).$$

Nach Definition von ε, (3), (13) und (14) ist jedes der k Glieder in $\mathfrak{P}$ für $|x| \leqq t$ absolut

$$< \varepsilon\,\mu^{1-N}\,P_2^{l_1+\cdots+l_{n-1}}(n)\,\omega^N(\tau) \leqq \varepsilon\,\omega^{1-N}(\tau)\,P_2^n(n)\,\omega^N(\tau) = \frac{\omega(\tau)}{2^{n+1}k},$$

also nach (4) für $|x| \leqq t$

$$y^{(n)} < k\frac{\omega(\tau)}{2^{n+1}k} + \frac{\omega(\tau)}{2^{n+1}} = \frac{\omega(\tau)}{2^n},$$

$$\underset{|x|\leqq t}{\text{obere Grenze von}}\ y^{(n)} \leqq \frac{\omega(\tau)}{2^n}.$$

Nach Definition von $\omega(t)$ ist daher für $0 \leqq t \leqq \frac{3}{4}$

$$\omega(t) \leqq \operatorname{Max}\left(4^n, \frac{\omega(\tau)}{2^n}\right),$$

(15)
$$\Omega(t) \leqq \operatorname{Max}\left(4, \frac{\Omega(\tau)}{2}\right).$$

Die Zahlen

$$\frac{\Omega\left(\frac{3}{4}-\frac{1}{2^\nu}\right)}{2^{\nu+1}}, \quad \nu \geqq 1 \text{ ganz},$$

beginnen mit $\dfrac{\Omega\left(\frac{1}{4}\right)}{4} \geqq 1$ und streben $\left(\text{da der Zähler } \leqq \Omega\left(\frac{3}{4}\right) \text{ ist}\right)$ bei $\nu \to \infty$ gegen 0. Für ein passendes $\nu \geqq 1$ ist also

(16)
$$\frac{\Omega\left(\frac{3}{4}-\frac{1}{2^\nu}\right)}{2^{\nu+1}} \geqq 1$$

und

(17)
$$\frac{\Omega\left(\frac{3}{4}-\frac{1}{2^{\nu+1}}\right)}{2^{\nu+2}} < 1.$$

Aus (16) und (15) folgt

(18)
$$2^{\nu+1} \leqq \Omega\left(\frac{3}{4}-\frac{1}{2^\nu}\right) \leqq \operatorname{Max}\left(4, \frac{1}{2}\Omega\left(\tau\left(\frac{3}{4}-\frac{1}{2^\nu}\right)\right)\right).$$

Hierin ist nach (16)

$$\tau\left(\frac{3}{4}-\frac{1}{2^\nu}\right) = \frac{3}{4} - \frac{1}{2^\nu} + \frac{1}{\Omega\left(\frac{3}{4}-\frac{1}{2^\nu}\right)} \leqq \frac{3}{4} - \frac{1}{2^\nu} + \frac{1}{2^{\nu+1}} = \frac{3}{4} - \frac{1}{2^{\nu+1}},$$

also nach (17)

$$\frac{1}{2}\Omega\left(\tau\left(\frac{3}{4}-\frac{1}{2^\nu}\right)\right) \leqq \frac{1}{2}\Omega\left(\frac{3}{4}-\frac{1}{2^{\nu+1}}\right) < 2^{\nu+1}.$$

Daher ist die rechte Seite von (18) gleich 4, also

$$\Omega\left(\frac{1}{4}\right) \leqq \Omega\left(\frac{3}{4} - \frac{1}{2^\nu}\right) \leqq 4.$$

Daher ist erstens

$$y^{(n)}(0) \leqq \mu \leqq \omega\left(\frac{1}{4}\right) \leqq 4^n \leqq P_2(n)\, 4^n = P(n),$$

zweitens nach (14) mit $t = 0$, $\tau = \frac{1}{\Omega(0)} \leqq \frac{1}{4}$

$$|y^{(m)}(0)| \leqq P_2(n)\, \Omega^m\left(\frac{1}{4}\right) \leqq P_2(n)\, 4^n = P(n) \quad \text{für} \quad 0 < m < n.$$

Göttingen, den 27. Februar 1929.

(Eingegangen am 2. 3. 1929.)

Über das asymptotische Verhalten normierter Lösungen von Differentialgleichungen mit Parameter.

Von

Wolfgang Sternberg in Breslau.

In einer in den Math. Annalen[1]) erschienenen Arbeit habe ich das asymptotische Verhalten der Lösungen gewisser elliptischer Differentialgleichungen mit Parameter für große reelle positive Werte dieses Parameters untersucht. Im § 4 dieser Arbeit handelte es sich insbesondere um „normierte" Lösungen der Differentialgleichung

$$\text{(A)} \qquad \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \varrho^2 w = 0.$$

Es zeigte sich, daß eine in einem gegebenen Bereiche S normierte, d. h. der Bedingung

$$\iint_S w^2\, dx\, dy = 1$$

genügende Lösung $w(x, y, \varrho)$ von (A) als Funktion von ϱ keineswegs für alle Punkte von S beschränkt zu sein braucht; es wurde nämlich eine normierte Lösung, welche in einem Punkte von S unendlich groß wie $\sqrt{\varrho}$ wird, tatsächlich angegeben. Dies war $w = \sqrt{\frac{\varrho}{2}}\, J_0(\varrho r)$, wo J_0 die Besselsche Funktion erster Art bezeichnete. Andrerseits ließ sich beweisen, daß die Ordnung des Unendlichwerdens der normierten Lösungen von (A) in bezug auf ϱ höchstens $\frac{1}{2}$ sein kann. Völlig analoge Sätze ergaben sich für die normierten Lösungen der Differentialgleichung

$$\text{(B)} \qquad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + [A(x, y, \varrho) + \varrho^2]\, u = 0,$$

wobei vorausgesetzt war, daß die Funktion A in S endliche Ableitungen erster Ordnung besitzt und nebst diesen Ableitungen bei wachsendem ϱ beschränkt bleibt.

---

[1]) 86 (1922), S. 280—295.

Daraus ergeben sich z. B. Konvergenzsätze für Reihen, die der Bilinearreihe in der Theorie der Integralgleichungen eng verwandt sind. Bedeutet nämlich $\varphi_n(x, y)$ die der Differentialgleichung (A) genügende, zum n-ten Eigenwert $\lambda_n = \varrho_n^2$ gehörige normierte Eigenfunktion des Bereiches S, wobei etwa als Randbedingung $\varphi_n = 0$ vorgeschrieben ist, so läßt sich jetzt leicht die Konvergenz der Reihe

$$\sum_{n=1}^{\infty} \frac{\varphi_n(x, y)\,\varphi_n(\xi, \eta)}{\lambda_n^{\frac{3}{2}+\varepsilon}}$$

für jedes positive ε beweisen. Da nämlich die Limesgleichung $\lim\limits_{n=\infty} \frac{\lambda_n}{n} = \text{konst.}$ (Weyl, Math. Annalen **71**, Crelles Journ. **141**, **143**; Courant, Math. Zeitschr. **7**) und andrerseits[2]) $\varphi_n(x, y) = O(\sqrt[4]{\lambda_n}) = O(\sqrt[4]{n})$ gilt, hat die obige Reihe die konvergente Majorantenreihe $\sum\limits_{n=1}^{\infty} \frac{1}{n^{1+\varepsilon}}$. Sie ist also für alle Punkte (x, y) und (ξ, η) von S gleichmäßig und absolut konvergent.

Ein ganz anderes Verhalten zeigen nun die normierten Lösungen der gewöhnlichen Differentialgleichung zweiter Ordnung

(1) $$y'' + \varrho^2 y = 0,$$

sowie auch der allgemeineren Sturm-Liouvilleschen Differentialgleichung

(2) $$z'' + [A(x) + \varrho^2] z = 0.$$

Für die Lösungen dieser Gleichungen kann man tatsächlich aus der Normiertheit auf die Beschränktheit schließen, was jetzt genauer ausgeführt werden soll.

Sei $y(x, \varrho)$ eine beliebige Lösung von (1), die im Intervalle $a \leq x \leq b$ zweimal stetig differentiierbar und normiert ist. Durch die lineare Transformation $x_1 = \frac{x-a}{b-a}$, durch die an der Form von (1) nichts Wesentliches geändert wird, kann man die Zurückführung auf das Intervall $0 \ldots 1$ erreichen, wodurch sich die Rechnung vereinfacht. Es wird also $a = 0$, $b = 1$ angenommen. Unsere Lösung kann in der Form

$$y = c_1(\varrho) \sin \varrho x + c_2(\varrho) \cos \varrho x$$

dargestellt werden, wo c_1 und c_2 Funktionen von ϱ allein sind. Nun ist

$$\int_0^1 y^2\,dx = c_1^2 \int_0^1 \sin^2 \varrho x\,dx + c_2^2 \int_0^1 \cos^2 \varrho x\,dx + 2 c_1 c_2 \int_0^1 \sin \varrho x \cos \varrho x\,dx$$
$$= 1.$$

[2]) Im folgenden wird das von E. Landau eingebürgerte Symbol O benutzt.

Oder

$$\int_0^1 y^2\,dx = c_1^2\left(\frac{1}{2} - \frac{\sin 2\varrho}{4\varrho}\right) + c_2^2\left(\frac{1}{2} + \frac{\sin 2\varrho}{4\varrho}\right) + 2c_1c_2\frac{1-\cos 2\varrho}{4\varrho}$$

$$= \frac{c_1^2+c_2^2}{2} + c_1^2\,O\left(\frac{1}{\varrho}\right) + c_2^2\,O\left(\frac{1}{\varrho}\right) + 2c_1c_2\,O\left(\frac{1}{\varrho}\right)$$

$$= \frac{c_1^2+c_2^2}{2} + (c_1^2+c_2^2)\,O\left(\frac{1}{\varrho}\right)$$

$$= (c_1^2+c_2^2)\left(\frac{1}{2} + O\left(\frac{1}{\varrho}\right)\right)$$

$$= 1.$$

Daher ist

$$c_1^2 + c_2^2 = O(1), \qquad c_1 = O(1), \qquad c_2 = O(1).$$

Mithin gilt im Intervall $0\ldots 1$ gleichmäßig die Limesgleichung

$$y(x,\varrho) = O(1),$$

w. z. b. w.

Was nun die Gleichung (2) betrifft, in der $A(x)$ für $0 \leqq x \leqq 1$ als stetig vorausgesetzt wird[8]), so benutzt man zuerst den fundamentalen Satz der Sturm-Liouvilleschen Theorie: Es existieren zwei für $0 \leqq x \leqq 1$ zweimal stetig differentiierbare Lösungen z_1, z_2 von (2), die den Limesgleichungen

$$z_1(x,\varrho) = \sin\varrho x + O\left(\frac{1}{\varrho}\right),$$

$$z_2(x,\varrho) = \cos\varrho x + O\left(\frac{1}{\varrho}\right)$$

genügen. Sie bilden natürlich ein Fundamentalsystem. Sei nun $z(x,\varrho)$ eine beliebige für $0 \leqq x \leqq 1$ zweimal stetig differentiierbare und normierte Lösung von (2).

Man kann

$$z = c_1(\varrho)z_1 + c_2(\varrho)z_2$$

setzen.

Da

$$\int_0^1 z_1^2\,dx = \frac{1}{2} + O\left(\frac{1}{\varrho}\right),$$

$$\int_0^1 z_2^2\,dx = \frac{1}{2} + O\left(\frac{1}{\varrho}\right),$$

$$\int_0^1 z_1z_2\,dx = O\left(\frac{1}{\varrho}\right)$$

[8]) Allgemeiner könnte man statt $A(x)$ eine Funktion $A(x,\varrho)$ zulassen, welche die Limesgleichung $A(x,\varrho) = O(1)$ erfüllt.

ist, findet man, wie oben,

$$\int_0^1 z^2\,dx = (c_1^2 + c_2^2)\left(\frac{1}{2} + O\left(\frac{1}{\varrho}\right)\right) = 1,$$

woraus sich wieder die Beschränktheit von c_1 und c_2 ergibt. Daraus folgt aber wegen $z_1 = O(1)$ und $z_2 = O(1)$ auch

$$z = O(1),$$

w. z. b. w.

Es sei noch bemerkt, daß ein analoger Satz für Gleichungen höherer Ordnung nicht gilt. So hat z. B. die Differentialgleichung vierter Ordnung

$$y^{\mathrm{IV}} - \varrho^4 y = 0$$

die in bezug auf das Intervall $0 \ldots 1$ normierte Lösung

$$y(x, \varrho) = \sqrt{2\varrho}\,\frac{e^{\varrho x}}{\sqrt{e^{2\varrho} - 1}}.$$

An der Stelle $x = 1$ hat man aber

$$y(1, \varrho) = \sqrt{2\varrho}\,\frac{e^{\varrho}}{\sqrt{e^{2\varrho} - 1}} = \sqrt{\varrho}\,O(1),$$

so daß $y(1, \varrho)$ unendlich groß und zwar von der Ordnung $\frac{1}{2}$ in bezug auf ϱ wird.

Ein ähnliches Resultat findet man bei Untersuchung der Gleichung

$$y^{\mathrm{IV}} + \varrho^4 y = 0.$$

Es ist also eine Eigenart der Sturm-Liouvilleschen Gleichung, daß sie keine normierten Lösungen besitzt, die nicht auch beschränkt sind. Hierin steht diese Gleichung in einem Gegensatz ebenso zu gewöhnlichen Differentialgleichungen höherer Ordnung wie auch zu partiellen Differentialgleichungen.

(Eingegangen am 1. 11. 1928.)

Theorie der adiabatischen Invarianten allgemeiner Differentialsysteme.

Von

Harald Geppert in Gießen, z. Zt. Rom. [1])

I.

Wie so viele mathematische Probleme verdankt auch die Frage nach den adiabatischen Invarianten mechanischer Systeme ihren Ursprung der Quantentheorie, und für ihre Zwecke wurde auch das erste auf sie bezügliche Theorem, der Sommerfeldsche Satz von der adiabatischen Invarianz der Phasenintegrale, bewiesen [2]). Daß diese Frage, unabhängig von den Strömungen der modernen Atomtheorie, auch für die Probleme der klassischen Mechanik von Wichtigkeit ist, hat man erst neuerdings erkannt, und man verdankt Herrn Levi-Civita [3]) eine sehr elegante, umfassende Theorie der adiabatischen Invarianten mechanischer Systeme, die die Basis für eine durchdringende mathematische Behandlung des ganzen Komplexes bildet. Die Levi-Civitasche Theorie unterliegt zwei wesentlichen Einschränkungen: Erstens gilt sie nur für Hamiltonsche Differentialsysteme, und zweitens setzt sie voraus, daß die bekannten Integrale derselben in Involution (im Lieschen Sinne) liegen, und es erhebt sich naturgemäß a posteriori die Frage, wie weit die erhaltenen Resultate von den genannten Bedingungen unabhängig sind. Bei der Untersuchung derselben zeigte sich, daß der Begriff der adiabatischen Invarianten viel weiter zu fassen ist als bisher und auf allgemeine Differentialsysteme übertragen werden muß, und daß sich alsdann eine weitreichende, in das Gebiet der Funktionalanalysis zu verweisende, mathematische Theorie desselben entwickeln läßt [4]), die reich an analytisch belangreichen Ergebnissen ist. Ihr sind die folgenden Zeilen gewidmet.

---

[1]) Fellow of the International Education Board.

[2]) Vgl. Burgers 1. 2. und die bei Born 1. wiedergegebenen Beweise.

[3]) Vgl. Levi-Civita 1. 2.

[4]) Vgl. die zusammenfassenden Noten: Geppert 1. 2. 3.

Ihrer Natur entsprechend hat die zu entwickelnde Theorie zwei Seiten, eine mathematische und eine physikalische. Wir werden in unserer Behandlung das Hauptgewicht auf die erstere legen, ohne den Kontakt mit den physikalisch wichtigen Problemen aufzugeben; dieser wird sich besonders in den die Mittelbildung betreffenden Fragen und der Klassifizierung der Systeme nach Imprimitivitätsordnungen bemerkbar machen. Zunächst setzen wir den allgemeinen Begriff der adiabatischen Invariante eines beliebigen Differentialsystems auseinander; diese Invarianten werden selbst durch ein neues Differentialsystem bestimmt, dessen nicht identisch befriedigte Integrabilitätsbedingungen die notwendigen und hinreichenden Voraussetzungen für die Existenz der adiabatischen Invarianten bilden. Es ist eine beachtenswerte Tatsache, daß sich für eine ausgedehnte Klasse von Fällen die Invarianten direkt angeben lassen, wodurch das Gibbs-Hertzsche Theorem verallgemeinert wird. Die entwickelte Theorie ist reich an Anwendungen auf die Mechanik auch nichthamiltonscher Systeme.

§ 1.

Der Begriff der adiabatischen Invariante.

Wir gehen aus von einem Differentialsystem erster Ordnung der Form:

$$\frac{dx_i}{dt} = X_i(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) \qquad (i = 1 \dots n) \tag{1.1}$$

in den Variablen $x_1 \dots x_n$. Die X_i seien (in dem zu betrachtenden Gebiete) stetig derivierbare Funktionen, die außer den genannten Variablen noch gewisse Parameter $a_1 \dots a_\varrho$ enthalten; wir setzen sie — ohne daß dies für die folgenden Begriffsbildungen wesentlich ist — als von t frei voraus. Sind die Parameter $a_1 \dots a_\varrho$ konstant, so besitzt das System Lösungen, die man formal schreiben kann:

$$x_i = x_i(t - t_0,\ c_1 \dots c_{n-1} \,|\, a_1 \dots a_\varrho) \qquad (i = 1 \dots n), \tag{1.2}$$

wobei $t_0, c_1 \dots c_{n-1}$ die n Integrationskonstanten bedeuten; sie sind im Kleinen eindeutig und daher in einem gewissen Bereich auflösbar in der Form:

$$\begin{aligned} f_i(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) &= c_i \qquad (i = 1 \dots n-1), \\ f_n(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) &= t - t_0 . \end{aligned} \tag{1.3}$$

Es wird zweckmäßig sein, sich diese Beziehungen im Raume R_n der $x_1 \dots x_n$ zu deuten; interpretieren wir t als Zeit, so stellen die Gleichungen (1.2) die Bahn eines sich im R_n bewegenden Punktes P dar, dessen Geschwindigkeitskomponenten durch (1.1) als Funktionen des Ortes gegeben sind. Wir wollen diese Trajektorie oder Integralkurve in Zukunft mit $\mathfrak{T}$

bezeichnen. Die ersten $n-1$ der Gleichungen (1.3) repräsentieren $n-1$ Hyperflächen des R_n der Dimension $n-1$, und $\mathfrak{T}$ erscheint als deren eindimensionale Schnittkurve. Dies alles gilt nur in einem Gebiete des R_n, in dem die Gleichungen (1.2) regulär und eindeutig auflösbar sind.

Sind nun die Parameter $a_1 \dots a_\varrho$ nicht mehr konstant, sondern auch Funktionen von t, so kann man zwar den Lösungen des Systems i. a. die Form (1.2) bzw. (1.3) belassen, muß dann aber auch die t_0, $c_1 \dots c_{n-1}$ als Funktionen von t ansehen, und zwar bestimmen sich letztere aus den auch der Störungstheorie zugrunde liegenden Differentialgleichungen:

$$(1.4)\qquad \frac{dc_i}{dt} = \sum_{\nu=1}^{\varrho} \frac{\partial f_i}{\partial a_\nu}\frac{da_\nu}{dt}, \qquad \frac{dt_0}{dt} = -\sum_{\nu=1}^{\varrho} \frac{\partial f_n}{\partial a_\nu}\frac{da_\nu}{dt} \qquad (i=1\dots n-1),$$

zu deren Integration wir zweckmäßig in $\frac{\partial f_i}{\partial a_\nu}$ $(i=1\dots n)$ die Relationen (1.2) an Stelle der x_i substituieren, wodurch wir etwa die Gleichungen:

$$(1.5)\qquad \frac{\partial f_i}{\partial a_\nu} = \varphi_{i\nu}(t-t_0,\ c_1\dots c_{n-1}\,|\,a_1\dots a_\varrho) \qquad (i=1\dots n;\ \nu=1\dots\varrho)$$

finden mögen. Das System (1.4) geht dann über in:

$$(1.6)\qquad dc_i = \sum_{\nu=1}^{\varrho} \varphi_{i\nu}\, da_\nu, \qquad dt_0 = -\sum_{\nu=1}^{\varrho} \varphi_{n\nu}\, da_\nu \qquad (i=1\dots n-1),$$

in dem nur noch die Variablen $t, t_0, c_1\dots c_{n-1}, a_1\dots a_\varrho$ vorkommen. Diese Gleichungen sind, wie aus (1.4) ersichtlich, bei vorgelegten $a_\nu(t)$ integrabel, und die $c_\nu(t)$ resultieren aus ihnen als *fonctions de ligne* bezüglich der $a_\nu(t)$; damit letztere zu gewöhnlichen Funktionen der a_ν werden, ist hingegen erforderlich, daß die Gleichungen (1.6) für beliebige a_ν unbeschränkt integrabel seien, was weitere Integrabilitätsbedingungen impliziert. Die Beziehungen (1.6) liefern zusammen mit (1.2) die Trajektorien der „gestörten" Bewegung. Man kann sie also, wenn man vom Falle konstanter Parameter als Grundlage ausgeht, als eine andere Form der Gleichungen (1.1) auffassen.

An Stelle dieses Differentialsystems (1.6), dessen Betrachtung nichts Neues geben würde, legen wir im folgenden ein anderes Differentialsystem zugrunde, das aus (1.6) dadurch hervorgeht, daß man die Koeffizienten $\varphi_{i\nu}$ einer Funktionaloperation bezüglich t unterwirft. Wir wollen hier speziell solche Operationen heranziehen, die den Charakter einer Mittelbildung tragen, jedoch ist auch das allgemeinere Problem einer beliebigen Operation der Behandlung zugänglich. Es sei also $T = T_0 \dots T_1$ ein beliebiges, aber festes Intervall der Variablen t, über das wir die Mittelbildung erstrecken, $F(t-t_0,\ c_1\dots c_{n-1}\,|\,a_1\dots a_\varrho)$ die Verteilungsfunktion der letzteren, so

setzen wir generell für eine von $t - t_0$, $c_1 \dots c_{n-1}$, $a_1 \dots a_\varrho$ abhängende Größe α als „Mittelwert“:

$$(1.7)\quad \bar{\alpha} = \int_{t_0+T_0}^{t_0+T_1} \alpha(t - t_0, c_1 \dots c_{n-1} | a_1 \dots a_\varrho) F(t - t_0, c_1 \dots c_{n-1} | a_1 \dots a_\varrho) dt : \int_{t_0+T_0}^{t_0+T_1} F(t - t_0, c_1 \dots c_{n-1} | a_1 \dots a_\varrho) dt,$$

natürlich unter der Voraussetzung, daß die auftretenden Integrale einen Sinn haben[5]). Die Integration in (1.7) verstehen wir so, daß sie sich nur auf das in α und F explizit auftretende Argument $t - t_0$ bezieht und demnach die Größen t_0, $c_1 \dots c_{n-1}$, $a_1 \dots a_\varrho$ als Konstanten behandelt.

Durch Anwendung der Operation (1.7) auf die Koeffizienten von (1.6) entsteht das neue n-gliedrige Differentialsystem:

$$(1.8)\quad dc_i = \sum_{\nu=1}^{\varrho} \bar{\varphi}_{i\nu} da_\nu \qquad dt_0 = -\sum_{\nu=1}^{\varrho} \bar{\varphi}_{n\nu} da_\nu \qquad (i = 1 \dots n-1),$$

das nur noch die Variablen t_0, $c_1 \dots c_{n-1}$, $a_1 \dots a_\varrho$ enthält, und in dem $a_1 \dots a_\varrho$ die Rolle der unabhängigen Variablen spielen. Es bildet die wesentliche Grundlage der folgenden Theorie. In das Gebiet der Physik verweisen wir die Frage, unter welchen Bedingungen das System (1.6) durch (1.8) *praktisch* ersetzt werden kann; es wird dies immer dann der Fall sein, wenn die Parameter sich gegenüber der Bewegung des Punktes P in einer zu präzisierenden Weise „unendlich langsam“ oder „adiabatisch“ ändern[6]). Wir setzen natürlich die Existenz der $\bar{\varphi}_{i\nu}$ voraus, wozu notwendig ist, daß T im Regularitätsintervall von (1.2) liege. Wir bezeichnen (1.8) kurz als das *gemittelte System.*

Das gemittelte System ist nicht immer komplett integrabel, wir werden vielmehr seine Integrabilitätsbedingungen, die sich in Einschränkungen für das ursprüngliche System (1.1) äußern, zu besprechen haben. Wir definieren nun:

Definition 1. *Jede Invariante des gemittelten Systems heißt eine adiabatische Invariante.*

Für eine adiabatische Invariante $J(t_0, c_1 \dots c_{n-1} | a_1 \dots a_\varrho)$ ist also charakteristisch das simultane Bestehen der folgenden ϱ Differentialgleichungen:

$$(1.9)\quad \sum_{i=1}^{n-1} \frac{\partial J}{\partial c_i} \bar{\varphi}_{i\nu} - \frac{\partial J}{\partial t_0} \bar{\varphi}_{n\nu} + \frac{\partial J}{\partial a_\nu} = 0 \qquad (\nu = 1 \dots \varrho).$$

---

[5]) In der Physik hat man bisher nur den Fall der einfachsten Mittelbildung: $F \equiv 1$ in Erwägung gezogen, doch bedeutet unser allgemeiner Standpunkt keine Erschwerung.

[6]) Vgl. hierzu die Note in Levi-Civita 3.

Die Systeme (1. 8) und (1. 9) sagen dasselbe aus und besitzen die gleichen Integrabilitätsbedingungen.

Nennen wir eine Größe „im Mittel invariant" gegenüber dem System (1. 1), wenn der durch (1. 7) erklärte Mittelwert ihrer totalen Ableitung nach t vermöge (1. 1) verschwindet, so können wir die obige Definition auch durch die folgende ersetzen:

Definition 2. *Eine für alle Werte der* $\frac{da_\nu}{dt}$ $(\nu = 1 \ldots \varrho)$ *im Mittel invariante Funktion der Integrationskonstanten und Parameter des Systems* (1. 1) *heißt eine adiabatische Invariante desselben.*

Man hat nämlich nur zu beachten, daß nach (1. 6)

$$\frac{dJ}{dt} = \sum_{\nu=1}^{\varrho} \left\{ \sum_{i=1}^{n-1} \frac{\partial J}{\partial c_i} \varphi_{i\nu} - \frac{\partial J}{\partial t_0} \varphi_{n\nu} + \frac{\partial J}{\partial a_\nu} \right\} \frac{da_\nu}{dt}$$

ist, und die Definition 2 fordert, da J von t nicht explizit abhängen soll, das Bestehen des Gleichungssystems (1. 9).

Diese zweite Definition ist die ursprünglich von den Physikern angenommene Begriffsbildung; denn für sie handelte es sich darum, Größen zu finden, die bei beliebiger adiabatischer Änderung der Parameter invariant bleiben, wobei das „unendlich schwach" der Änderung dadurch präzisiert wird, daß an die Stelle der gewöhnlichen Invarianz die Invarianz im Mittel treten darf. Dies hat praktisch den Vorteil, daß man im System (1. 9) zu Gleichungen gelangt, die viel leichter zu integrieren sind als die ursprünglichen Beziehungen (1. 6); wir werden, wie sich noch zeigen wird, in der Lage sein, in einer großen Klasse von Fällen ihre Integrale wirklich anzugeben.

Die durch (1. 7) definierten Mittelwerte hängen i. a. noch von t_0 ab; wir wollen uns aber im folgenden ausschließlich auf solche Fälle beschränken, in denen dies nicht eintritt, nämlich in denen das Intervall T nicht von t_0 abhängt, und unser Interesse in erster Linie auf die Fälle konzentrieren, in denen die Mittelwerte über unendliches Intervall erstreckt werden, also

$$(1.10)\quad \bar{\alpha} = \lim_{T \to \infty} \int_{t_0}^{t_0+T} \alpha(t - t_0, c_1 \ldots c_{n-1} \,|\, a_1 \ldots a_\varrho) F(t - t_0, c_1 \ldots c_{n-1} \,|\, a_1 \ldots a_\varrho)\, dt$$
$$: \int_{t_0}^{t_0+T} F(t - t_0, c_1 \ldots c_{n-1} \,|\, a_1 \ldots a_\varrho)\, dt$$

zu setzen ist, weil hier unsere Theorie besonders schöne Resultate liefern wird. Man kann dann das Problem der adiabatischen Invarianten etwas enger fassen und sich auf die Aufsuchung solcher Invarianten beschränken, die von t_0 unabhängig sind, d. h. solcher Funktionen $J(c_1 \ldots c_{n-1} \,|\, a_1 \ldots a_\varrho)$,

für die die Gleichungen

$$\sum_{i=1}^{n-1} \frac{\partial J}{\partial c_i} \bar{\varphi}_{i\nu} + \frac{\partial J}{\partial a_\nu} = 0 \qquad (\nu = 1 \dots \varrho) \tag{1.11}$$

identisch befriedigt sind.

Da jede Funktion einer oder mehrerer adiabatischen Invarianten wieder eine solche ist, wird es genügen, die unabhängigen partikulären Lösungen von (1.11) zu finden, deren größtmögliche Anzahl $n-1$ sein wird; sie werden die Größen $c_1 \dots c_{n-1}$ mit den Parametern in Beziehung setzen. Die Gleichungen (1.11) setzen die Existenz der $\bar{\varphi}_{i\nu}$ und diese wieder die eindeutige Auflösbarkeit von (1.2) im Gebiete T voraus; sollte diese Bedingung für eine Reihe von Größen $\bar{\varphi}_{i\nu}$ nicht erfüllt sein, so kann trotzdem das Problem seinen Sinn behalten, indem man J von den entsprechenden c_i als unabhängig voraussetzt. Schließlich entnimmt man aus (1.11) den trivialen Satz:

Satz 1. *Jedes von den Parametern freie Integral $f_i(x_1 \dots x_n) = c_i$ des ursprünglichen Systems ist eine adiabatische Invariante.*

§ 2.

Zwei Beispiele.

Bevor wir uns der allgemeinen Theorie zuwenden, mögen hier zwei Beispiele Platz finden, die wir später von andern Gesichtspunkten aus behandeln werden.

1. Beispiel. Die eindimensionale elastische Bewegung kann, wenn a_1 die Elastizitätskonstante bedeutet, folgendermaßen dargestellt werden:

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -a_1 x_1.$$

Die Integrale sind:

$$f_1 \equiv \sqrt{x_1^2 + \frac{1}{a_1} x_2^2} = c_1, \qquad f_2 \equiv \operatorname{arctang} \sqrt{a_1} \frac{x_1}{x_2} = t - t_0.$$

Nimmt man $F \equiv 1$, $T = \infty$ oder, was auf dasselbe hinauskommt, gleich der Periode $\frac{2\pi}{\sqrt{a_1}}$ der Bewegung, so wird

$$\bar{\varphi}_{11} = -\frac{c_1}{4a_1},$$

und (1.11) lautet:

$$\frac{\partial J}{\partial a_1} - \frac{c_1}{4a_1} \frac{\partial J}{\partial c_1} = 0,$$

deren allgemeine Lösung

$$J = f(a_1 c_1^4) \tag{2.1}$$

ist; d. h. bei adiabatischer Variation der Elastizitätskonstanten ändert sich die Amplitude der Schwingung umgekehrt proportional der vierten Wurzel aus der Elastizitätskonstanten.

2. Beispiel. Das Differentialsystem

$$x_1 \frac{dx_1}{dt} = a_1^2 x_2, \qquad \frac{dx_2}{dt} = -a_2^2$$

besitzt die Integrale

$$f_1 \equiv x_1^2 + \frac{a_1^2}{a_2^2} x_2^2 = c_1, \qquad f_2 \equiv -\frac{x_2}{a_2^2} = t - t_0,$$

und die (elliptische) Trajektorie ist nur dann reell, wenn $t - t_0$ zwischen den Grenzen $-\frac{\sqrt{c_1}}{a_1 a_2}$ und $\frac{\sqrt{c_1}}{a_1 a_2}$ liegt. Nehmen wir dieses Intervall als T und $F \equiv 1$, so lautet das System (1.9) folgendermaßen:

$$\frac{\partial J}{\partial a_1} + \frac{2}{3} \frac{c_1}{a_1} \frac{\partial J}{\partial c_1} = 0, \qquad \frac{\partial J}{\partial a_2} - \frac{2}{3} \frac{c_1}{a_2} \frac{\partial J}{\partial c_1} = 0,$$

es ist komplett integrabel und besitzt die Lösung

$$J = f\left(\frac{a_2}{a_1} c_1^{3/2};\ t_0\right) \tag{2.2}$$

als adiabatische Invariante.

Auf das später zu behandelnde Beispiel des gedämpften Oszillators wollen wir hier nur verweisen.

§ 3.

Zweidimensionale Systeme.

Indem wir uns nunmehr dem allgemeinen Problem zuwenden, werden wir vorerst den Fall der zweidimensionalen Systeme ($n = 2$) erledigen, weil er sich vollkommen behandeln läßt und einen Leitfaden für die folgenden Betrachtungen abgibt. Es handelt sich also um das System

$$\frac{dx_1}{dt} = X_1(x_1, x_2 \mid a_1 \ldots a_\varrho); \qquad \frac{dx_2}{dt} = X_2(x_1, x_2 \mid a_1 \ldots a_\varrho), \tag{3.1}$$

dessen Integral (wir lassen den Index 1 weg) lautet:

$$f(x_1, x_2 \mid a_1 \ldots a_\varrho) = c, \tag{3.2}$$

und das mithin die Gleichung

$$\frac{\partial f}{\partial x_1} X_1 + \frac{\partial f}{\partial x_2} X_2 \equiv 0 \tag{3.3}$$

befriedigen muß. Man entnimmt dieser Beziehung die Existenz einer Größe μ derart, daß

$$\frac{\partial f}{\partial x_1} = -\mu X_2, \qquad \frac{\partial f}{\partial x_2} = \mu X_1 \tag{3.4}$$

ist. Die Integrabilitätsbedingung von (3.4) erfordert, daß μ eine Lösung der Gleichung

$$(3.5) \qquad \frac{\partial}{\partial x_1}(\mu X_1) + \frac{\partial}{\partial x_2}(\mu X_2) \equiv X_1 \frac{\partial \mu}{\partial x_1} + X_2 \frac{\partial \mu}{\partial x_2} + \mu\left(\frac{\partial X_1}{\partial x_1} + \frac{\partial X_2}{\partial x_2}\right) = 0,$$

d. h. daß μ ein sogenannter Jacobischer Multiplikator des Systems (3.1) sei. Vorerst ist μ jedoch nicht ein beliebiger Multiplikator, sondern eben derjenige, der durch die Gleichungen (3.4) eindeutig festgelegt ist; μ ist eine Funktion von $x_1, x_2 \mid a_1 \dots a_\varrho$.

Wir wollen annehmen, daß die Trajektorie $\mathfrak{T}$ frei von allen Singularitäten sei, derart, daß im besonderen die beiden Ableitungen $\frac{\partial f}{\partial x_1}$ und $\frac{\partial f}{\partial x_2}$ nicht gleichzeitig verschwinden; dies besagt dann, daß μ zufolge (3.4) längs $\mathfrak{T}$ und damit wegen der Stetigkeit auch in einer gewissen Umgebung von $\mathfrak{T}$ niemals verschwindet, wenn wir annehmen, daß X_1, X_2 daselbst sich regulär verhalten. μ hat also in seinem Regularitätsgebiet längs $\mathfrak{T}$ ein bestimmtes Vorzeichen. Das Bogenelement von $\mathfrak{T}$ läßt sich dann wegen (3.1) und (3.4) auf die Form bringen:

$$(3.6) \quad ds = \sqrt{dx_1^2 + dx_2^2} = dt \cdot \sqrt{X_1^2 + X_2^2} = \frac{dt}{\mu} \cdot \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2} = dt \cdot \frac{\Gamma_{12}}{\mu},$$

wobei wir gesetzt haben:

$$(3.7) \qquad \Gamma_{12}^2 = \left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2.$$

Unseren Voraussetzungen gemäß verschwindet auch Γ_{12} längs $\mathfrak{T}$ niemals und kann — indem der Richtungssinn von s nach fortschreitendem t orientiert wird — als mit μ von gleichem Vorzeichen angenommen werden.

Diese Beziehung (3.6) versetzt uns in die Lage, die bisher in unsern Definitionen auftretenden Mittelwerte nach t durch räumliche, über $\mathfrak{T}$ zu erstreckende Integralbildungen zu ersetzen. Bei diesem Übergang tritt dann an Stelle des Intervalls T der Variablen $t - t_0$ vermöge (1.2) ein räumliches Intervall S auf $\mathfrak{T}$ für die Variable s; ferner muß die bisherige Verteilungsfunktion $F(t - t_0, c_1 \mid a_1 \dots a_\varrho)$ vermöge (1.3) durch eine eindeutige reguläre räumliche Verteilungsfunktion $F(x_1, x_2 \mid a_1 \dots a_\varrho)$ ersetzt werden, so daß die zugrunde gelegte Operation (1.7) übergeht in

$$(3.8) \quad \bar{u} = \int_S u(x_1, x_2 \mid a_1 \dots a_\varrho)\, F(x_1, x_2 \mid a_1 \dots a_\varrho) \frac{\mu}{\Gamma_{12}}\, ds : \int_S F(x_1, x_2 \mid a_1 \dots a_\varrho) \frac{\mu}{\Gamma_{12}}\, ds.$$

Hier ist eine Bemerkung über μ einzufügen; μ war bisher ein durch (3.4) eindeutig bestimmter Multiplikator von (3.1). Nun gilt aber der leicht zu beweisende Satz, daß der Quotient zweier verschiedener Multiplikatoren eine Invariante des ursprünglichen Systems ist, mit andern Worten: auf

der Trajektorie $\mathfrak{T}$ unterscheiden sich die Multiplikatoren nur um einen konstanten Faktor. Da letzterer in (3.8) sich forthebt, darf daselbst nunmehr μ jeden beliebigen Multiplikator, d. h. jede Lösung von (3.5) bedeuten.

Ist das System (3.1) ein Liouvillesches System, d. h. verschwindet seine „Divergenz“:

$$\frac{\partial X_1}{\partial x_1} + \frac{\partial X_2}{\partial x_2} = 0, \tag{3.9}$$

so zeigt (3.5), daß $\mu =$ konst. gesetzt und demnach μ aus (3.8) und allen folgenden Beziehungen gestrichen werden kann. Ein Spezialfall der Liouvilleschen sind wiederum die Hamiltonschen Systeme [7]).

Infolge von (3.8) nehmen die Bestimmungsgleichungen (1.11) der adiabatischen Invarianten die Form an:

$$\frac{\partial J}{\partial c} \cdot \int\limits_S \frac{\partial f}{\partial a_\nu} F \frac{\mu}{\Gamma_{12}} ds + \frac{\partial J}{\partial a_\nu} \cdot \int\limits_S F \frac{\mu}{\Gamma_{12}} ds = 0 \qquad (\nu = 1 \ldots \varrho). \tag{3.10}$$

Der wichtigste zu betrachtende Fall ist der der einfachen Mittelbildung $F \equiv 1$. Bevor wir an die Lösung dieser Gleichungen gehen, müssen wir ihre Integrabilitätsbedingungen untersuchen.

§ 4.

Die notwendigen Integrabilitätsbedingungen.

Die Bedingungen für die vollständige Integrabilität des Systems (1.11) oder, was auf dasselbe hinauskommt, der ersten $n-1$ Gleichungen (1.8) lauten in unserm zweidimensionalen Falle:

$$\frac{d}{da_\lambda}\bar\varphi_{1\varkappa} - \frac{d}{da_\varkappa}\bar\varphi_{1\lambda} \equiv \frac{\partial}{\partial a_\lambda}\bar\varphi_{1\varkappa} - \frac{\partial}{\partial a_\varkappa}\bar\varphi_{1\lambda} + \bar\varphi_{1\lambda}\frac{\partial}{\partial c}\bar\varphi_{1\varkappa} - \bar\varphi_{1\varkappa}\frac{\partial}{\partial c}\bar\varphi_{1\lambda} = 0 \tag{4.1}$$
$$(\varkappa, \lambda = 1 \ldots \varrho).$$

Führen wir vorübergehend die Bezeichnung ein:

$$g = \int\limits_S F \frac{\mu}{\Gamma_{12}} ds; \qquad g_\nu = \int\limits_S \frac{\partial f}{\partial a_\nu} F \frac{\mu}{\Gamma_{12}} ds \qquad (\nu = 1 \ldots \varrho), \tag{4.2}$$

so kann man (4.1) in die Form setzen:

$$g\left(\frac{\partial g_\varkappa}{\partial a_\lambda} - \frac{\partial g_\lambda}{\partial a_\varkappa}\right) + g_\lambda\left(\frac{\partial g_\varkappa}{\partial c} + \frac{\partial g}{\partial a_\varkappa}\right) - g_\varkappa\left(\frac{\partial g_\lambda}{\partial c} + \frac{\partial g}{\partial a_\lambda}\right) = 0 \tag{4.3}$$
$$(\varkappa, \lambda = 1 \ldots \varrho),$$

und es ist unsere Aufgabe, die hierin auftretenden Derivierten wirklich zu berechnen.

[7]) Eine andere Ableitung von (3.8) für letztere findet sich bei Levi-Civita 1, S. 335–337.

Das zu betrachtende Stück S der Trajektorie $\mathfrak{T}$ möge sich von dem Punkte P_0, der dem Werte $t - t_0 = T_0$ entspreche, bis zu P_1, der für $t - t_0 = T_1$ angenommen wird, erstrecken, derart, daß $T_1 - T_0 = T$ das Mittelungsintervall bedeutet. Erstreckt sich die Mittelung über die gesamte unendlich lange Trajektorie, so ist im folgenden die Betrachtung der Randpunkte P_0, P_1 überflüssig. Sowohl T_0 als T_1 sind Funktionen von $c \mid a_1 \ldots a_\varrho$. Es wird dann zweckmäßig sein, neben der Kurvenschar (3.2) der Trajektorien $\mathfrak{T}$ die der sogenannten Synchronen (vgl. (1.3)):

$$f_2(x_1, x_2 \mid a_1 \ldots a_\varrho) = t - t_0 = \text{konst.}, \tag{4.4}$$

die wir generell mit $\mathfrak{S}$ bezeichnen wollen, einzuführen und jeden der zu betrachtenden Punkte als Schnitt einer $\mathfrak{T}$ mit (mindestens) einer $\mathfrak{S}$ aufzufassen. Die Randpunkte werden auf $\mathfrak{T}$ von den Synchronen

$$f_2(x_1, x_2 \mid a_1 \ldots a_\varrho) = T_0 \text{ bzw. } T_1 \tag{4.5}$$

ausgeschnitten. Ändern wir nun c allein um dc und lassen $a_1 \ldots a_\varrho$ fest, so geht $\mathfrak{T}$ in eine neue Trajektorie $\mathfrak{T}'$ mit der Gleichung

$$f(x_1, x_2 \mid a_1 \ldots a_\varrho) = c + dc \tag{4.6}$$

über, und deren Randpunkte P_0', P_1' werden von den Synchronen

$$f_2(x_1, x_2 \mid a_1 \ldots a_\varrho) = T_0 + \frac{\partial T_0}{\partial c} dc \quad \text{bzw.} \quad T_1 + \frac{\partial T_1}{\partial c} dc \tag{4.7}$$

ausgeschnitten.

Wir ordnen jedem Punkte P von $\mathfrak{T}$ den auf seiner Normalen n liegenden Schnittpunkt mit $\mathfrak{T}'$ zu und nennen diesen P'; diese Zuordnung läßt den Punkten P_0 bzw. P_1 zwei Punkte $\overline{P}_0'$ bzw. $\overline{P}_1'$ entsprechen, die im allgemeinen von P_0', P_1' verschieden sein werden. Um die in (4.3) auftretenden Derivierten nach c zu berechnen, müssen wir die Integrale $g_\varkappa$ längs $\mathfrak{T}'$ von P_0' bis P_1' erstrecken und den längs $\mathfrak{T}$ genommenen Wert derselben subtrahieren. Wir können also schreiben:

$$\frac{\partial g_\varkappa}{\partial c} = \lim_{dc \to 0} \frac{1}{dc} \left\{ \int_{P_0'}^{\overline{P}_0'} + \int_{\overline{P}_1'}^{P_1'} + \left(\int_{\overline{P}_0'}^{\overline{P}_1'} - \int_{P_0}^{P_1} \right) \right\}; \tag{4.8}$$

die beiden ersten Summanden nennen wir die (infinitesimalen) Randglieder, der dritte läßt sich vermöge unserer normalen Zuordnung $P \to P'$ leicht berechnen. Es sei nämlich $d'n$ die normale Verschiebung $P' - P$ und ihre Komponenten $d'x_1$, $d'x_2$, also

$$d'x_1 = \frac{1}{\Gamma_{12}} \frac{\partial f}{\partial x_1} d'n, \qquad d'x_2 = \frac{1}{\Gamma_{12}} \frac{\partial f}{\partial x_2} d'n, \tag{4.9}$$

und es mögen durch Akzente alle auf $\mathfrak{T}'$ bezüglichen Größen bezeichnet werden; dann ist bis auf Größen zweiter Ordnung:

$$d's = \sqrt{d(x_1+d'x_1)^2+d(x_2+d'x_2)^2} = ds\left\{1+\frac{1}{\Gamma_{12}}\frac{\partial f}{\partial x_2}\frac{d}{ds}(d'x_1)-\frac{1}{\Gamma_{12}}\frac{\partial f}{\partial x_1}\frac{d}{ds}(d'x_2)\right\}$$

$$(4.10)\qquad = ds\left\{1+\frac{d'n}{\Gamma_{12}^3}\left[\frac{\partial^2 f}{\partial x_1^2}\left(\frac{\partial f}{\partial x_2}\right)^2-2\frac{\partial^2 f}{\partial x_1\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}+\frac{\partial^2 f}{\partial x_2^2}\left(\frac{\partial f}{\partial x_1}\right)^2\right]\right\},$$

$$\frac{1}{\Gamma_{12}'}=\frac{1}{\Gamma_{12}}\left\{1-\frac{d'n}{\Gamma_{12}^3}\left[\frac{\partial^2 f}{\partial x_1^2}\left(\frac{\partial f}{\partial x_1}\right)^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}+\frac{\partial^2 f}{\partial x_2^2}\left(\frac{\partial f}{\partial x_2}\right)^2\right]\right\}.$$

Aus (4.6) und (4.9) folgt aber andrerseits:

$$\frac{\partial f}{\partial x_1}d'x_1+\frac{\partial f}{\partial x_2}d'x_2 = \Gamma_{12}\,d'n = dc,$$

somit

$$(4.11)\qquad d'n=\frac{dc}{\Gamma_{12}};\qquad d'x_1=\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}dc,\qquad d'x_2=\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}dc,$$

und damit sind wir in der Lage, den dritten Summanden in (4.8) hinzuschreiben.

Wir berechnen noch die Randglieder. Die Koordinaten von $\overline{P}_0'$ sind nach (4.11)

$$\overline{P}_0':\qquad x_1+\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}dc\Big|^0,\qquad x_2+\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}dc\Big|^0,$$

diejenigen von P_0' dagegen finden sich aus den beiden Gleichungen (4.6) und (4.7). Bedeutet nämlich $(\delta'x_1, \delta'x_2)$ den Vektor P_0P_0', so ist in P_0

$$\frac{\partial f}{\partial x_1}\delta'x_1+\frac{\partial f}{\partial x_2}\delta'x_2=dc,$$

$$\frac{\partial f_2}{\partial x_1}\delta'x_1+\frac{\partial f_2}{\partial x_2}\delta'x_2=\frac{\partial T_0}{\partial c}dc,$$

also, wenn wir vorübergehend

$$\mathfrak{F}=\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_2}-\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_1}$$

setzen:

$$(4.12)\qquad \delta'x_1=\frac{dc}{\mathfrak{F}}\left(\frac{\partial f_2}{\partial x_2}-\frac{\partial f}{\partial x_2}\frac{\partial T_0}{\partial c}\right)\Big|^0,\qquad \delta'x_2=\frac{dc}{\mathfrak{F}}\left(-\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_1}\frac{\partial T_0}{\partial c}\right)\Big|^0.$$

Aus den wegen (3.1) folgenden Gleichungen

$$(4.13)\qquad \frac{\partial f}{\partial x_1}X_1+\frac{\partial f}{\partial x_2}X_2=0,\qquad \frac{\partial f_2}{\partial x_1}X_1+\frac{\partial f_2}{\partial x_2}X_2=1$$

folgt, daß $\mathfrak{F}\neq 0$ ist. Die Komponenten von $P_0'\overline{P}_0'$ sind demnach:

$$dc\left\{\frac{1}{\mathfrak{F}}\left(\frac{\partial f_2}{\partial x_2}-\frac{\partial f}{\partial x_2}\frac{\partial T_0}{\partial c}\right)-\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}\right\}\Big|^0;\qquad dc\left\{\frac{1}{\mathfrak{F}}\left(-\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_1}\frac{\partial T_0}{\partial c}\right)-\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}\right\}\Big|^0$$

und daher die Länge dieser Strecke:

$$\overline{P_0'\overline{P}_0'}=\frac{dc}{\mathfrak{F}\Gamma_{12}}\left[\Gamma_{12}^2\frac{\partial T_0}{\partial c}-\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Big|^0$$

$$=dc\cdot\sqrt{X_1^2+X_2^2}\left[\frac{\partial T_0}{\partial c}-\frac{1}{\Gamma_{12}^2}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Big|^0.$$

Ein entsprechender Ausdruck findet sich für den andern Randpunkt; die Randglieder in der Summe (4.8) haben also schließlich den Wert:

$$\frac{\partial f}{\partial a_\varkappa} F\mu \frac{\sqrt{X_1^2+X_2^2}}{\Gamma_{12}}\left[\frac{\partial T_{0,1}}{\partial c} - \frac{1}{\Gamma_{12}^2}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Bigg|_{P_1}^{P_0}.$$

Damit sind wir endlich in der Lage, den Ausdruck für $\frac{\partial g_\varkappa}{\partial c}$ zusammenzusetzen; man hat nach Vorangehendem:

$$(4.14)\left\{\begin{aligned}\frac{\partial g_\varkappa}{\partial c} = &-\int_S \frac{\partial f}{\partial a_\varkappa} F \frac{\mu}{\Gamma_{12}^5}\left\{\left(\frac{\partial^2 f}{\partial x_1^2} - \frac{\partial^2 f}{\partial x_2^2}\right)\left(\left(\frac{\partial f}{\partial x_1}\right)^2 - \left(\frac{\partial f}{\partial x_2}\right)^2\right) + 4\frac{\partial^2 f}{\partial x_1 \partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\right\} ds\\ &+\int_S F\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial^2 f}{\partial a_\varkappa \partial x_1}\frac{\partial f}{\partial x_1} + \frac{\partial^2 f}{\partial a_\varkappa \partial x_2}\frac{\partial f}{\partial x_2}\right) ds\\ &+\int_S \frac{\partial f}{\partial a_\varkappa} F \frac{1}{\Gamma_{12}^3}\left(\frac{\partial \mu}{\partial x_1}\frac{\partial f}{\partial x_1} + \frac{\partial \mu}{\partial x_2}\frac{\partial f}{\partial x_2}\right) ds\\ &+\int_S \frac{\partial f}{\partial a_\varkappa}\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial F}{\partial x_1}\frac{\partial f}{\partial x_1} + \frac{\partial F}{\partial x_2}\frac{\partial f}{\partial x_2}\right) ds\\ &+\frac{\partial f}{\partial a_\varkappa} F\mu \frac{\sqrt{X_1^2+X_2^2}}{\Gamma_{12}}\left[\frac{\partial T_{0,1}}{\partial c} - \frac{1}{\Gamma_{12}^2}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Bigg|_{P_1}^{P_0}.\end{aligned}\right.$$

In gleicher Weise leisten wir nun die Berechnung der Derivierten von g und $g_\varkappa$ nach a_λ. Geht a_λ in $a_\lambda + da_\lambda$ über, während alle übrigen Größen festbleiben, so tritt an Stelle von $\mathfrak{T}$ die Kurve $\mathfrak{T}^*$ mit der Gleichung

$$(4.15)\qquad f(x_1, x_2 \mid a_1 \ldots a_\lambda + da_\lambda \ldots a_\varrho) = c,$$

und die Randpunkte werden auf ihr durch die Synchronen

$$(4.16)\qquad f_2(x_1, x_2 \mid a_1 \ldots a_\lambda + da_\lambda \ldots a_\varrho) = T_0 + \frac{\partial T_0}{\partial a_\lambda} da_\lambda \quad \text{bzw.} \quad T_1 + \frac{\partial T_1}{\partial a_\lambda} da_\lambda$$

ausgeschnitten; wir bezeichnen sie mit P_0^*, P_1^*, während $\overline{P}_0^*, \overline{P}_1^*$ die normal über P_0, P_1 auf $\mathfrak{T}^*$ liegenden Punkte bezeichnen sollen. Jedem Punkte P von $\mathfrak{T}$ ordnen wir den auf seiner Normalen liegenden Punkt P^* zu und bezeichnen mit d^*n den Vektor PP^*, dessen Komponenten

$$(4.17)\qquad d^*x_1 = \frac{1}{\Gamma_{12}}\frac{\partial f}{\partial x_1} d^*n, \qquad d^*x_2 = \frac{1}{\Gamma_{12}}\frac{\partial f}{\partial x_2} d^*n$$

sind. Eine zu (4.8) entsprechende Zerlegung führt auch hier zum Ziel. Nun folgt aber aus (4.15) und (4.17)

$$\frac{\partial f}{\partial x_1} d^*x_1 + \frac{\partial f}{\partial x_2} d^*x_2 \equiv \Gamma_{12} d^*n = -\frac{\partial f}{\partial a_\lambda} da_\lambda,$$

also

$$(4.18)\qquad d^*n = -\frac{\partial f}{\partial a_\lambda}\frac{da_\lambda}{\Gamma_{12}};$$

$$d^*x_1 = -\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial a_\lambda} da_\lambda, \qquad d^*x_2 = -\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}\frac{\partial f}{\partial a_\lambda} da_\lambda.$$

Die auf $\mathfrak{T}^*$ bezüglichen Größen erhalten im folgenden einen Stern; dann ist nach (4.10), worin nur d^*n statt $d'n$ zu substituieren ist:

$$d^*s = ds\left\{1 - \frac{da_\lambda}{\Gamma_{12}^4}\frac{\partial f}{\partial a_\lambda}\left[\frac{\partial^2 f}{\partial x_1^2}\left(\frac{\partial f}{\partial x_2}\right)^2 - 2\frac{\partial^2 f}{\partial x_1\,\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2} + \frac{\partial^2 f}{\partial x_2^2}\left(\frac{\partial f}{\partial x_1}\right)^2\right]\right\}; \tag{4.19}$$

in gleicher Weise findet man aus der auf (4.10) folgenden Formel, wenn man beachtet, daß a_λ in f auch explizit auftritt:

$$\begin{aligned}\frac{1}{\Gamma_{12}^*} = \frac{1}{\Gamma_{12}}\Bigg\{1 &+ \frac{da_\lambda}{\Gamma_{12}^4}\frac{\partial f}{\partial a_\lambda}\left[\frac{\partial^2 f}{\partial x_1^2}\left(\frac{\partial f}{\partial x_1}\right)^2 + 2\frac{\partial^2 f}{\partial x_1\,\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2} + \frac{\partial^2 f}{\partial x_2^2}\left(\frac{\partial f}{\partial x_2}\right)^2\right]\\ &- \frac{da_\lambda}{\Gamma_{12}^2}\left[\frac{\partial f}{\partial x_1}\frac{\partial^2 f}{\partial x_1\,\partial a_\lambda} + \frac{\partial f}{\partial x_2}\frac{\partial^2 f}{\partial x_2\,\partial a_\lambda}\right]\Bigg\},\end{aligned} \tag{4.20}$$

und mittels dieser Formeln sind wir in der Lage, die dem dritten Summanden von (4.8) entsprechenden Terme hinzuschreiben.

Es bleiben wieder die Randglieder zu bestimmen. Wegen (4.18) hat $\overline{P}_0^*$ die Koordinaten

$$\overline{P}_0^*: \qquad x_1 - \frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial a_\lambda}da_\lambda\Big|^0, \quad x_2 - \frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}\frac{\partial f}{\partial a_\lambda}da_\lambda\Big|^0.$$

P_0^* hingegen möge die Koordinaten $x_i + \delta^* x_i|^0$ haben, dann folgt aus (4.15) und (4.16):

$$\frac{\partial f}{\partial x_1}\delta^* x_1 + \frac{\partial f}{\partial x_2}\delta^* x_2 = -\frac{\partial f}{\partial a_\lambda}da_\lambda,$$

$$\frac{\partial f_2}{\partial x_1}\delta^* x_1 + \frac{\partial f_2}{\partial x_2}\delta^* x_2 = \frac{\partial(T_0 - f_2)}{\partial a_\lambda}da_\lambda,$$

gültig in P_0, und demnach:

$$\begin{aligned}\delta^* x_1 &= \frac{da_\lambda}{\mathfrak{F}}\left(-\frac{\partial f_2}{\partial x_2}\frac{\partial f}{\partial a_\lambda} - \frac{\partial f}{\partial x_2}\frac{\partial(T_0 - f_2)}{\partial a_\lambda}\right),\\ \delta^* x_2 &= \frac{da_\lambda}{\mathfrak{F}}\left(+\frac{\partial f_2}{\partial x_1}\frac{\partial f}{\partial a_\lambda} + \frac{\partial f}{\partial x_1}\frac{\partial(T - f_2)}{\partial a_\lambda}\right).\end{aligned} \tag{4.21}$$

Die Komponenten des Vektors $P_0^*\overline{P}_0^*$ sind also

$$da_\lambda\left[-\frac{1}{\mathfrak{F}}\left(\frac{\partial f_2}{\partial x_2}\frac{\partial f}{\partial a_\lambda} + \frac{\partial f}{\partial x_2}\frac{\partial(T_0 - f_2)}{\partial a_\lambda}\right) + \frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial a_\lambda}\right],$$

$$da_\lambda\left[\ \frac{1}{\mathfrak{F}}\left(\frac{\partial f_2}{\partial x_1}\frac{\partial f}{\partial a_\lambda} + \frac{\partial f}{\partial x_1}\frac{\partial(T_0 - f_2)}{\partial a_\lambda}\right) + \frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial x_2}\frac{\partial f}{\partial a_\lambda}\right],$$

und daher die Länge dieses Vektors:

$$\begin{aligned}\overline{P_0^*\overline{P}_0^*} &= \frac{da_\lambda}{\mathfrak{F}\Gamma_{12}}\left[\frac{\partial(T_0 - f_2)}{\partial a_\lambda}\Gamma_{12}^2 + \frac{\partial f}{\partial a_\lambda}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Big|^0\\ &= da_\lambda\cdot\sqrt{X_1^2 + X_2^2}\left[\frac{\partial(T_0 - f_2)}{\partial a_\lambda} + \frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial a_\lambda}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Big|^0,\end{aligned}$$

und ein entsprechender Ausdruck gilt im andern Endpunkt. Damit lauten beispielsweise die den beiden ersten Summanden in (4.8) entsprechenden Randglieder in der Zerlegung für $\frac{\partial g}{\partial a_\lambda}$:

$$F\mu\frac{\sqrt{X_1^2+X_2^2}}{\Gamma_{12}}\left[\frac{\partial(T_{0,1}-f_2)}{\partial a_\lambda}+\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial a_\lambda}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Bigg|_{P_1}^{P_0}$$

und analog für $\frac{\partial g_\varkappa}{\partial a_\lambda}$, wo nur der Faktor $\frac{\partial f}{\partial a_\varkappa}$ hinzutritt.

Setzt man alles zusammen, so findet man endlich:

$$(4.22)\left\{\begin{aligned}\frac{\partial g}{\partial a_\lambda} = &\int\limits_S \frac{\partial f}{\partial a_\lambda}F\frac{\mu}{\Gamma_{12}^5}\left\{\left(\frac{\partial^2 f}{\partial x_1^2}-\frac{\partial^2 f}{\partial x_2^2}\right)\left(\left(\frac{\partial f}{\partial x_1}\right)^2-\left(\frac{\partial f}{\partial x_2}\right)^2\right)+4\frac{\partial^2 f}{\partial x_1\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\right\}ds\\ &-\int\limits_S F\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial^2 f}{\partial a_\lambda\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial a_\lambda\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\lambda}F\frac{1}{\Gamma_{12}^3}\left(\frac{\partial\mu}{\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial\mu}{\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\lambda}\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial F}{\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial F}{\partial x_2}\frac{\partial f}{\partial x_2}\right)ds+\int\limits_S\left(F\frac{\partial\mu}{\partial a_\lambda}+\mu\frac{\partial F}{\partial a_\lambda}\right)\frac{1}{\Gamma_{12}}ds\\ &+F\mu\frac{\sqrt{X_1^2+X_2^2}}{\Gamma_{12}}\left[\frac{\partial(T_{0,1}-f_2)}{\partial a_\lambda}+\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial a_\lambda}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Bigg|_{P_1}^{P_0},\end{aligned}\right.$$

und:

$$(4.23)\left\{\begin{aligned}\frac{\partial g_\varkappa}{\partial a_\lambda} = &\int\limits_S \frac{\partial f}{\partial a_\varkappa}\frac{\partial f}{\partial a_\lambda}F\frac{\mu}{\Gamma_{12}^5}\left\{\left(\frac{\partial^2 f}{\partial x_1^2}-\frac{\partial^2 f}{\partial x_2^2}\right)\left(\left(\frac{\partial f}{\partial x_1}\right)^2-\left(\frac{\partial f}{\partial x_2}\right)^2\right)+4\frac{\partial^2 f}{\partial x_1\partial x_2}\frac{\partial f}{\partial x_1}\frac{\partial f}{\partial x_2}\right\}ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\varkappa}F\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial^2 f}{\partial a_\lambda\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial a_\lambda\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\lambda}F\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial^2 f}{\partial a_\varkappa\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial^2 f}{\partial a_\varkappa\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\varkappa}\frac{\partial f}{\partial a_\lambda}F\frac{1}{\Gamma_{12}^3}\left(\frac{\partial\mu}{\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial\mu}{\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &-\int\limits_S \frac{\partial f}{\partial a_\varkappa}\frac{\partial f}{\partial a_\lambda}\frac{\mu}{\Gamma_{12}^3}\left(\frac{\partial F}{\partial x_1}\frac{\partial f}{\partial x_1}+\frac{\partial F}{\partial x_2}\frac{\partial f}{\partial x_2}\right)ds\\ &+\int\frac{\partial f}{\partial a_\varkappa}\left(F\frac{\partial\mu}{\partial a_\lambda}+\mu\frac{\partial F}{\partial a_\lambda}\right)\frac{1}{\Gamma_{12}}ds+\int\limits_S\frac{\partial^2 f}{\partial a_\varkappa\partial a_\lambda}F\frac{\mu}{\Gamma_{12}}ds\\ &+\frac{\partial f}{\partial a_\varkappa}F\mu\frac{\sqrt{X_1^2+X_2^2}}{\Gamma_{12}}\left[\frac{\partial(T_{0,1}-f_2)}{\partial a_\lambda}+\frac{1}{\Gamma_{12}^2}\frac{\partial f}{\partial a_\lambda}\left(\frac{\partial f}{\partial x_1}\frac{\partial f_2}{\partial x_1}+\frac{\partial f}{\partial x_2}\frac{\partial f_2}{\partial x_2}\right)\right]\Bigg|_{P_1}^{P_0}.\end{aligned}\right.$$

Diese etwas länglichen Formeln versetzen uns in die Lage, die Integrabilitätsbedingungen (4. 3) in extenso hinzuschreiben. Man findet zunächst aus (4. 14), (4. 22) und (4. 23) die einfachen Gleichungen:

$$
\begin{aligned}
(4.24)\qquad \frac{\partial g_\varkappa}{\partial a_\lambda} - \frac{\partial g_\lambda}{\partial a_\varkappa} &= \int\limits_S \left\{ \frac{\partial f}{\partial a_\varkappa} \frac{\partial(\mu F)}{\partial a_\lambda} - \frac{\partial f}{\partial a_\lambda} \frac{\partial(\mu F)}{\partial a_\varkappa} \right\} \frac{1}{\Gamma_{12}}\, ds \\
&\quad + F\mu \frac{\sqrt{X_1^2 + X_2^2}}{\Gamma_{12}} \left\{ \frac{\partial f}{\partial a_\varkappa} \frac{\partial(T_{0,1} - f_2)}{\partial a_\lambda} - \frac{\partial f}{\partial a_\lambda} \frac{\partial(T_{0,1} - f_2)}{\partial a_\varkappa} \right\} \Bigg|_{P_1}^{P_0},
\end{aligned}
$$

$$
(4.25)\qquad \frac{\partial g_\varkappa}{\partial c} + \frac{\partial g}{\partial a_\varkappa} = \int\limits_S \frac{\partial(\mu F)}{\partial a_\varkappa} \frac{1}{\Gamma_{12}}\, ds + F\mu \frac{\sqrt{X_1^2 + X_2^2}}{\Gamma_{12}} \left\{ \frac{\partial(T_{0,1} - f_2)}{\partial a_\varkappa} + \frac{\partial f}{\partial a_\varkappa} \frac{\partial T_{0,1}}{\partial c} \right\} \Bigg|_{P_1}^{P_0}.
$$

Wir wollen uns im folgenden ausschließlich auf den Fall beschränken, daß in diesen beiden Gleichungen die auf P_0 und P_1 bezüglichen Randglieder für sich verschwinden, was eine jetzt zu interpretierende Bedingung für T_0, T_1 darstellt. Das von uns geforderte Verschwinden ist offenbar identisch damit, daß in den Endpunkten P_0, P_1 die Beziehungen

$$
(4.26)\qquad \frac{\partial(f_2 - T_{0,1})}{\partial a_\varkappa} = \frac{\partial f}{\partial a_\varkappa} \frac{\partial T_{0,1}}{\partial c} \qquad (\varkappa = 1 \ldots \varrho)
$$

gelten. Aus (4. 12) und (4. 21) entnimmt man dann, daß daselbst

$$
(4.27)\qquad \frac{\delta' x_1}{\delta' x_2} = \frac{\delta^* x_1}{\delta^* x_2}
$$

wird, d. h. daß dann die Punkte P_0, P_0' und P_0^* bzw. P_1, P_1', P_1^* auf einer Geraden liegen. Umgekehrt, liegen diese Punktetripel je auf einer Geraden, so muß (4. 27) gelten, und aus (4. 12), (4. 21) entnimmt man, da, wie schon oben bemerkt, $\mathfrak{F} \neq 0$ ist, das Bestehen von (4. 26). Man kann das auch anders ausdrücken: Nehmen wir nämlich an, die Randpunkte lägen je auf einer von den $a_1 \ldots a_\varrho$ unabhängigen Kurve $\mathfrak{C}$ mit der Gleichung

$$
(4.28)\qquad \Pi(x_1, x_2) = 0,
$$

so müssen alle Randpunkte mit gleichem Index auf der Tangente an $\mathfrak{C}$ in P_0 bzw. P_1 liegen, indem nämlich aus (4. 28) folgt

$$
\frac{\delta' x_1}{\delta' x_2} = \frac{\delta^* x_1}{\delta^* x_2} = - \frac{\partial \Pi}{\partial x_2} : \frac{\partial \Pi}{\partial x_1} \Bigg|^{0,1},
$$

es muß also (4. 26) gelten; und umgekehrt, wenn die Randpunkte auf einer Geraden liegen, so werden sie von einer Kurve, deren Gleichung die $a_1 \ldots a_\varrho$ nicht enthält (z. B. eben jener Geraden) ausgeschnitten. Unsere Forderung bezüglich des Verschwindens der Randterme ist also gleichbedeutend mit der Annahme:

Forderung: ***Das Mittelungsintervall S erstreckt sich entweder über die unendliche Länge der Trajektorien, oder wird auf diesen von einer festen, die Parameter nicht enthaltenden Kurve ausgeschnitten.***

Wir haben dabei die *beiden* Kurven, die P_0 und P_1 ausschneiden, in eine einzige zusammengefaßt, was stets erlaubt ist.

Ist dieses Postulat erfüllt, so geben (4.3), (4.24) und (4.25) schließlich die Integrabilitätsbedingungen:

$$\begin{aligned}(4.29)\quad &\int\limits_S F\frac{\mu}{\Gamma_{12}}ds\cdot\int\limits_S\left\{\frac{\partial f}{\partial a_\varkappa}\frac{\partial(\mu F)}{\partial a_\lambda}-\frac{\partial f}{\partial a_\lambda}\frac{\partial(\mu F)}{\partial a_\varkappa}\right\}\frac{1}{\Gamma_{12}}ds+\int\limits_S\frac{\partial f}{\partial a_\lambda}F\frac{\mu}{\Gamma_{12}}ds\cdot\int\limits_S\frac{\partial(\mu F)}{\partial a_\varkappa}\frac{1}{\Gamma_{12}}ds\\ &+\int\limits_S\frac{\partial f}{\partial a_\varkappa}F\frac{\mu}{\Gamma_{12}}ds\cdot\int\limits_S\frac{\partial(\mu F)}{\partial a_\lambda}\frac{1}{\Gamma_{12}}ds=0\qquad(\varkappa,\lambda=1\dots\varrho).\end{aligned}$$

Verwenden wir die Bezeichnung (3.8), so können wir ihnen folgende, sehr übersichtliche Form geben:

$$(4.30)\quad \overline{\frac{\partial f}{\partial a_\varkappa}\frac{\partial\log\mu F}{\partial a_\lambda}}-\overline{\frac{\partial f}{\partial a_\varkappa}}\cdot\overline{\frac{\partial\log\mu F}{\partial a_\lambda}}=\overline{\frac{\partial f}{\partial a_\lambda}\frac{\partial\log\mu F}{\partial a_\varkappa}}-\overline{\frac{\partial f}{\partial a_\lambda}}\cdot\overline{\frac{\partial\log\mu F}{\partial a_\varkappa}}\quad(\varkappa,\lambda=1\dots\varrho),$$

in der sie eine gewisse Permutabilität von Mittelbildung und Multiplikation fordern.

Wir beschließen diesen Paragraphen mit dem

Satz 2. *Die notwendigen und hinreichenden Bedingungen für die Existenz der adiabatischen Invarianten eines zweidimensionalen Systems lauten* (4.30).

§ 5.

Hinreichende Existenzbedingungen.

Wir wenden uns nun der Deutung und Verwertung des Gefundenen zu. Längs $\mathfrak{T}$ können wir die Größen $\frac{\partial f}{\partial a_\varkappa}$, $\frac{\partial\log\mu F}{\partial a_\varkappa}$ als Funktionen des Parameters $t-t_0$ ausdrücken; setzen wir vorübergehend:

$$\overline{\frac{\partial f}{\partial a_\varkappa}}=\alpha_\varkappa,\quad \overline{\frac{\partial\log\mu F}{\partial a_\varkappa}}=\beta_\varkappa\qquad(\varkappa=1\dots\varrho),$$

so können wir eine Zerlegung vornehmen:

$$(5.1)\qquad \frac{\partial f}{\partial a_\varkappa}=\alpha_\varkappa+h_\varkappa(t-t_0),\quad \frac{\partial\log\mu F}{\partial a_\varkappa}=\beta_\varkappa+j_\varkappa(t-t_0),$$

worin $h_\varkappa$, $j_\varkappa$ Funktionen ihres Arguments bezeichnen, deren Mittelwert über das Intervall T genommen verschwindet. Die Bedingung (4.30) erhält dann die Form:

$$(5.2)\qquad \overline{h_\varkappa j_\lambda}=\overline{h_\lambda j_\varkappa}\qquad(\varkappa,\lambda=1\dots\varrho).$$

Was sie bedeutet, kann man sich etwa an dem Falle einer periodischen Lösung, in dem $h_\varkappa$ und $j_\varkappa$ zu Fourierreihen werden, klarmachen; doch wollen wir sie nicht in ihrer Allgemeinheit behandeln, sondern drei Spezialfälle herausgreifen, die der weiteren Analyse zugänglich sind:

Fall I. Es sind alle

$$h_\varkappa(t-t_0)=0 \qquad (\varkappa=1\ldots\varrho),$$

d. h. auf $\mathfrak{T}$ ist $\frac{\partial f}{\partial a_\varkappa}$ konstant, es muß also

$$\frac{\partial f}{\partial a_\varkappa}=\Phi_\varkappa(f|a_1\ldots a_\varrho) \qquad (\varkappa=1\ldots\varrho) \tag{5.3}$$

sein. Dann folgen neben (3.3) durch Differentiation nach $a_\varkappa$ noch die Gleichungen

$$\frac{\partial f}{\partial x_1}\frac{\partial X_1}{\partial a_\varkappa}+\frac{\partial f}{\partial x_2}\frac{\partial X_2}{\partial a_\varkappa}=0 \qquad (\varkappa=1\ldots\varrho), \tag{5.4}$$

d. h. f ist nicht nur Integral des ursprünglichen Systems (3.1), sondern auch der ϱ folgenden:

$$\frac{dx_1}{dt}=\frac{\partial X_1}{\partial a_\varkappa},\quad \frac{dx_2}{dt}=\frac{\partial X_2}{\partial a_\varkappa} \qquad (\varkappa=1\ldots\varrho), \tag{5.5}$$

was besagt, daß f gegenüber den Parametern ein stationäres Integral ist. In den Gleichungen (1.8) verschwindet die Funktionaloperation, sie werden mit (1.6), d. h. hier mit

$$dc=\sum_{\nu=1}^{\varrho}\frac{\partial f}{\partial a_\nu}da_\nu$$

identisch. Weiterhin ergeben (3.3) und (5.4) die Proportionen:

$$X_1:X_2=\frac{\partial X_1}{\partial a_\varkappa}:\frac{\partial X_2}{\partial a_\varkappa} \qquad (\varkappa=1\ldots\varrho),$$

$$\frac{\partial}{\partial a_\varkappa}\left(\frac{X_1}{X_2}\right)=0,$$

woraus folgt, daß auf der Trajektorie $\mathfrak{T}$ der Richtungsfaktor

$$\frac{dx_1}{dx_2}=\frac{X_1}{X_2}$$

eine Funktion von x_1, x_2 allein ist. Man kann mithin die Kurvenschar der $\mathfrak{T}$ von den Parametern unabhängig wählen und ein Integral $f(x_1, x_2)$ von (3.1) finden, das die Parameter überhaupt nicht enthält und für das dann nach Satz 1 c die adiabatische Invariante wird. Dieser Fall ist also erledigt.

Fall II. Für einen bestimmten Multiplikator μ ist in (5.1)

$$j_\varkappa(t-t_0)=0 \qquad (\varkappa=1\ldots\varrho),$$

also längs $\mathfrak{T}$ ist $\frac{\partial \log \mu F}{\partial a_\varkappa}$ konstant, mithin:

$$\frac{\partial \log \mu F}{\partial a_\varkappa} = \Psi_\varkappa(f|a_1 \ldots a_\varrho) \qquad (\varkappa = 1 \ldots \varrho). \tag{5.6}$$

Diese Bedingung ist im besonderen dann erfüllt, wenn die ϱ Gleichungen

$$\frac{\partial \log \mu F}{\partial a_\varkappa} = 0 \qquad (\varkappa = 1 \ldots \varrho) \tag{5.7}$$

gelten, also μF von den Parametern gar nicht abhängt — ein Fall, mit dem wir uns gleich eingehender beschäftigen werden. Die Beziehungen (5.6) lassen eine einfache Deutung zu, wenn die Verteilungsfunktion $F \equiv 1$ ist. Es bezeichne nämlich

$$\Delta_0 = \frac{\partial X_1}{\partial x_1} + \frac{\partial X_2}{\partial x_2} \tag{5.8}$$

die Divergenz des Systems (3.1), dann kann man (3.5) schreiben:

$$\frac{\partial \log \mu}{\partial x_1} X_1 + \frac{\partial \log \mu}{\partial x_2} X_2 + \Delta_0 = 0, \tag{5.9}$$

und die Differentiation nach $a_\varkappa$ gibt, wenn (5.6) mit $F \equiv 1$ besteht:

$$\frac{\partial \log \mu}{\partial x_1} \frac{\partial X_1}{\partial a_\varkappa} + \frac{\partial \log \mu}{\partial x_2} \frac{\partial X_2}{\partial a_\varkappa} + \frac{\partial \Delta_0}{\partial a_\varkappa} = 0 \qquad (\varkappa = 1 \ldots \varrho), \tag{5.10}$$

was besagt, daß μ auch ein Multiplikator der Systeme (5.5), oder mit andern Worten ein gegenüber den Parametern stationärer Multiplikator ist. Aus (5.9) und (5.10) folgt, daß in der Matrix

$$\begin{pmatrix} X_1 & X_2 & \Delta_0 \\ \frac{\partial X_1}{\partial a_\varkappa} & \frac{\partial X_2}{\partial a_\varkappa} & \frac{\partial \Delta_0}{\partial a_\varkappa} \end{pmatrix} \qquad (\varkappa = 1 \ldots \varrho)$$

sämtliche dreireihigen Determinanten verschwinden müssen. (5.7) ist speziell erfüllt, wenn $F \equiv 1$ und $\mu = \text{konst.}$, also das System (3.1) ein Liouvillesches ist. Damit ist als Spezialfall unserer Untersuchungen der auf Gibbs und Hertz[8]) zurückgehende Satz bewiesen:

Satz 3. *Ein zweidimensionales Hamiltonsches Differentialsystem besitzt bei einfacher Mittelbildung eine adiabatische Invariante.*

Fall III. $\qquad j_\varkappa(t - t_0) = \text{konst.}\, h_\varkappa(t - t_0) \qquad (\varkappa = 1 \ldots \varrho),$

wobei die Konstante für alle $\varkappa$ die gleiche ist. Aus (5.1) folgt dann, daß

$$\frac{\partial \log \mu F}{\partial a_\varkappa} = p(f|a_1 \ldots a_\varrho) \frac{\partial f}{\partial a_\varkappa} + q_\varkappa(f|a_1 \ldots a_\varrho) \qquad (\varkappa = 1 \ldots \varrho) \tag{5.11}$$

---

[8]) Vgl. Gibbs 1., Hertz 1., S. 534—535.

sein muß. Bestimmt man dann ein $\varphi(f|a_1\ldots a_\varrho)$ derart, daß

$$\frac{\partial \log\varphi(f|a_1\ldots a_\varrho)}{\partial f} = p(f|a_1\ldots a_\varrho)$$

ist, so gibt (5. 11)

$$\frac{\partial \log(\mu F:\varphi)}{\partial a_\varkappa} = -\frac{1}{\varphi}\frac{\partial\varphi}{\partial a_\varkappa} + q_\varkappa = \Psi_\varkappa(f|a_1\ldots a_\varrho) \qquad (\varkappa = 1\ldots\varrho),$$

und da mit μ zugleich auch $\mu:\varphi$ ein Multiplikator von (3. 1) ist, kommen wir auf den Fall II zurück.

Wir fassen unsere Resultate zusammen:

Satz 4. *Hinreichende Existenzbedingungen für die adiabatische Invariante eines zweidimensionalen Systems sind* a) *die Stationarität eines Integrals,* b) *bei einfacher Mittelbildung die Stationarität eines Multiplikators gegenüber den Parametern.*

§ 6.

Die Invariante der generalisiert-Liouvilleschen Systeme.

Bereits in (5. 7) haben wir den Fall hervorgehoben, in dem ein Multiplikator μ angegeben werden kann, derart, daß μF von den Parametern $a_1\ldots a_\varrho$ unabhängig ist; wir bezeichnen dann das System (3. 1) aus einem später ersichtlichen Grunde als generalisiert-Liouvillesches System. Für dasselbe muß eine adiabatische Invariante existieren. Die Gleichungen (4. 24), (4. 25) vereinfachen sich (unter Beibehaltung der Forderung des § 4) zu:

$$(6.1) \qquad \frac{\partial g_\varkappa}{\partial a_\lambda} = \frac{\partial g_\lambda}{\partial a_\varkappa}, \qquad \frac{\partial g}{\partial a_\varkappa} + \frac{\partial g_\varkappa}{\partial c} = 0 \qquad (\varkappa, \lambda = 1\ldots\varrho),$$

und es muß sich demnach eine Funktion $V(c|a_1\ldots a_\varrho)$ finden lassen derart, daß

$$(6.2) \qquad \frac{\partial V}{\partial c} = g, \qquad \frac{\partial V}{\partial a_\varkappa} = -g_\varkappa$$

ist; die Gleichungen (3. 10) zeigen dann, daß V die gesuchte adiabatische Invariante unseres zweidimensionalen Systems wird.

Wir können V leicht explizit angeben. Bezeichnet nämlich $\mathfrak{C}$ eine feste, von den $a_1\ldots a_\varrho$ unabhängige Kurve, die gemäß unserer Forderung das Mittelungsintervall S auf $\mathfrak{T}$ ausschneidet, bzw. eine beliebige feste Kurve, wenn $T = \infty$ ist, und ist B der von $\mathfrak{C}$ und S begrenzte Bereich der x_1, x_2-Ebene, so ist

$$(6.3) \qquad J = V = \iint_B \mu F\, dx_1\, dx_2,$$

vorausgesetzt, daß dieses Integral einen Sinn hat. Der Beweis, daß (6. 2) erfüllt ist, ist mittels der Entwicklungen des § 4 leicht zu führen.

Bei bloßer Variation von c bleiben nämlich μF und $\mathfrak{C}$ ungeändert, nur die Trajektorie $\mathfrak{T}$ geht in $\mathfrak{T}'$ über; die normale Verschiebung der Punkte P in P' wird dabei durch (4.11) charakterisiert, und es wird also

$$\frac{\partial V}{\partial c} = \int\limits_S \mu F\, ds\, d'n : dc = \int\limits_S F \frac{\mu}{\Gamma_{12}}\, ds = g;$$

man überzeugt sich leicht, daß die Unterscheidung von P_0', $\overline{P}_0'$ bzw. P_1', $\overline{P}_1'$ für die Berechnung dieser Derivierten belanglos ist, indem die Dreiecke $P_0 P_0' \overline{P}_0'$ bzw. $P_1 P_1' \overline{P}_1'$ unendlich klein von zweiter Ordnung sind ($\mathfrak{C}$ kann ja offenbar $\mathfrak{T}$ nicht in P_0, P_1 berühren, da sonst nach (4.12) $\mathfrak{F} = 0$ sein müßte). Bei bloßer Variation von $a_\varkappa$ bleibt nach Voraussetzung (5.7) μF ebenso wie $\mathfrak{C}$ unbeeinflußt, und $\mathfrak{T}$ geht durch die normale Verschiebung $d^* n$ in $\mathfrak{T}^*$ über, so daß mittels (4.18) folgt:

$$\frac{\partial V}{\partial a_\varkappa} = \int\limits_S \mu F\, ds\, d^* n : da_\varkappa = -\int\limits_S \frac{\partial f}{\partial a_\varkappa} F \frac{\mu}{\Gamma_{12}}\, ds = -g_\varkappa,$$

wie wir behauptet hatten. Wir haben also folgenden Satz gefunden:

Satz 5. *Die adiabatische Invariante eines zweidimensionalen generalisiert-Liouvilleschen Systems ist* $V = \iint\limits_B \mu F\, dx_1\, dx_2$.

Ist $\mathfrak{T}$ im besonderen periodisch und $T = \infty$ oder gleich der Periode, so wird man für B den von $\mathfrak{T}$ eingeschlossenen Bereich nehmen können. Handelt es sich um einfache Mittelbildung: $F = 1$, so fallen die Liouvilleschen Systeme ($\Delta_0 = 0$, $\mu =$ konst.) unter die betrachtete Klasse; sie lassen sich im zweidimensionalen Falle (aber auch nur hier) in die Klasse der Hamiltonschen Systeme transformieren[9]), und Satz 5 fällt dann mit der Aussage des Gibbs-Hertzschen Theorems[10]) zusammen, daß das „Phasenvolumen"

$$V = \iint dx_1\, dx_2$$

adiabatisch invariant ist. Die Gleichung (5.7) verlangt dann, daß μ von den Parametern unabhängig sei, und für solche Systeme gilt ein Satz, den wir später gleich für n Dimensionen beweisen und hier einstweilen nur aussprechen wollen:

Satz 6. *Ist μ von den Parametern unabhängig, so gibt es eine von den Parametern freie Koordinatentransformation, die das ursprüngliche System in ein Liouvillesches überführt, und umgekehrt.*

Solche, die Parameter nicht enthaltende Transformationen sind natürlich gestattet und für die Berechnung der adiabatischen Invarianten ohne

---

[9]) Vgl. Levi-Civita 1., S. 336.

[10]) Vgl. Gibbs 1., Hertz 1., S. 534–535, Levi-Civita 1., S. 339–342.

Einfluß, und man sieht daher, daß die genannte Klasse von Systemen im wesentlichen mit der der Liouvilleschen Systeme zusammenfällt, und dadurch erklärt sich die am Anfang des Paragraphen eingeführte Bezeichnung.

§ 7.

Beispiele. Der gedämpfte Oszillator.

Die beiden Beispiele des § 2 lassen sich mittels der vorangehenden Sätze sofort erledigen.

1. Beispiel. Es ist $\Delta_0 = 0$, $\mu = \text{konst.} = 1$; als B nehmen wir wegen der Periodizität das Innere der Ellipse

$$x_1^2 + \frac{1}{a_1} x_2^2 = c_1^2,$$

so daß

$$J = V = \iint_B dx_1\, dx_2 = \pi \sqrt{a_1}\, c^2$$

wird, was mit (2.1) übereinstimmt.

2. Beispiel. Es ist

$$\Delta_0 = -a_1^2 \frac{x_2}{x_1^2}, \qquad \mu = x_1,$$

wir haben also den Fall des § 6 vor uns. Demgemäß existiert eine adiabatische Invariante. Als Kurve $\mathfrak{C}$, die die Endpunkte bestimmt, nehmen wir die Achse $x_1 = 0$ und demnach als B den rechten (oder linken), von der x_2-Achse begrenzten Teil der Ellipse

$$x_1^2 + \frac{a_1^2}{a_2^2} x_2^2 = c_1,$$

so daß also nach Satz 5 folgt:

$$J = \iint_B x_1\, dx_1\, dx_2 = \frac{2}{3} \frac{a_2}{a_1} c_1^{3/2},$$

was mit (2.2) in Einklang steht.

3. Beispiel. Schließlich wollen wir, weil von physikalischem Interesse, den gedämpften Oszillator behandeln; der harmonische Oszillator fällt unter die klassischen Beispiele[11]). Es bezeichnen a_1 und a_2 die Frequenz- bzw. Dämpfungskonstante des Oszillators; dessen Bewegung wird dann durch die Differentialgleichungen:

$$\frac{dx_1}{dt} = x_2, \qquad \frac{dx_2}{dt} = -2a_2 x_2 - a_1^2 x_1 \tag{7.1}$$

[11]) Vgl. Born 1., Levi-Civita 1., S. 345; 2., S. 13–14; wir verdanken den Hinweis auf dieses Beispiel Herrn Levi-Civita.

wiedergegeben, deren Integrale lauten:

$$(7.2)\quad \begin{aligned} x_1 &= c\,e^{-a_2(t-t_0)}\cos\gamma(t-t_0)\\ x_2 &= -c\,e^{-a_2(t-t_0)}\{a_2\cos\gamma(t-t_0)+\gamma\sin\gamma(t-t_0)\}, \end{aligned}$$

worin gesetzt ist:

$$\gamma = \sqrt{a_1^2 - a_2^2}.$$

In aufgelöster Form lauten diese Gleichungen:

$$f(x_1, x_2\,|\,a_1, a_2) \equiv \frac{1}{\gamma}\sqrt{\gamma^2 x_1^2 + (a_2 x_1 + x_2)^2}\; e^{-\frac{a_2}{\gamma}\arctang\frac{a_2 x_1 + x_2}{\gamma x_1}} = c,$$

$$f_2(x_1, x_2\,|\,a_1, a_2) \equiv -\frac{1}{\gamma}\arctang\frac{a_2 x_1 + x_2}{\gamma x_1} = t - t_0,$$

so daß man findet:

$$(7.3)\left\{\begin{aligned} \frac{\partial f}{\partial a_1} &= \varphi_{11} = -\frac{a_1 c}{\gamma^2}\left\{a_2(t-t_0)+\sin^2\gamma(t-t_0)+\frac{a_2}{2\gamma}\sin 2\gamma(t-t_0)\right\},\\ \frac{\partial f}{\partial a_2} &= \varphi_{12} = \frac{c}{2\gamma^3}\{2a_1^2\gamma(t-t_0)+(2a_2^2-a_1^2)\sin 2\gamma(t-t_0)\\ &\qquad\qquad - 2a_2\gamma\cos 2\gamma(t-t_0)\},\\ \frac{\partial f_2}{\partial a_1} &= \varphi_{21} = -\frac{a_1}{\gamma^2}\left\{(t-t_0)+\frac{1}{2\gamma}\sin 2\gamma(t-t_0)\right\},\\ \frac{\partial f_2}{\partial a_2} &= \varphi_{22} = \frac{1}{\gamma^2}\left\{a_2(t-t_0)-\cos^2\gamma(t-t_0)+\frac{a_2}{2\gamma}\sin 2\gamma(t-t_0)\right\}. \end{aligned}\right.$$

Die Trajektorie ist eine sich immer enger um den Nullpunkt wickelnde Spirale; als die Endpunkte des Mittelungsintervalls bestimmende Kurve $\mathfrak{C}$ nehmen wir die x_2-Achse, wählen

$$T_0 = \frac{\pi}{2\gamma},\quad T_1 = \frac{(2n+1)\pi}{2\gamma},\quad T = \frac{n\pi}{\gamma}$$

und finden demnach die Mittelwerte ($F \equiv 1$ angenommen):

$$\bar{\varphi}_{11} = -\frac{a_1 c}{2\gamma^3}\{a_2\pi(n+1)+\gamma\},\quad \bar{\varphi}_{12} = \frac{a_1^2 c(n+1)\pi}{2\gamma^3},$$

so daß sich die adiabatische Invariante aus den beiden Gleichungen (1.11), d. h.

$$\frac{\partial J}{\partial c}\cdot a_1 c\{a_2\pi(n+1)+\gamma\} - \frac{\partial J}{\partial a_1}\cdot 2\gamma^3 = 0,$$

$$\frac{\partial J}{\partial c}\cdot a_1^2 c(n+1)\pi \qquad + \frac{\partial J}{\partial a_2}\cdot 2\gamma^3 = 0$$

bestimmen müßte. Man sieht aber sofort, daß die Integrabilitätsbedingung (4.1) hier nicht erfüllt ist, indem deren linke Seite den Wert

$$-\frac{a_1 a_2 c}{\gamma^4}$$

annimmt. Den gleichen Wert ergibt die linke Seite von (4.29), wenn man beachtet, daß die Größe

$$\mu = e^{-\frac{2a_2}{\gamma}\operatorname{arctang}\frac{a_2 x_1 + x_2}{\gamma x_1}}$$

ein Multiplikator des Systems (7.1) ist.

Es gibt also keine adiabatische Invariante im absoluten Sinne, d. h. keine Größe, die bei ***beliebiger*** adiabatischer Änderung der a_1, a_2 invariant bliebe; womit übrigens der Nachweis erbracht ist, daß unsere Existenzbedingungen (4.30) nicht trivial sind. Das gleiche gilt a fortiori für $T = \infty$, weil da die auftretenden Mittelwerte überhaupt nicht existieren.

Aber eine andere Fragestellung führt zu bemerkenswerten Ergebnissen; wir fragen: Kann man eine solche Funktion $\lambda(a_1, a_2)$ angeben, daß ***bei absoluter Konstanz von*** λ eine adiabatische Invariante existiert? Dann sind also die adiabatischen Änderungen von a_1, a_2 nicht mehr beliebig, sondern durch die Konstanz von λ beschränkt, und eine solche Invariante trägt den Charakter einer ***relativen adiabatischen Invariante.*** Wir werden also statt a_1, a_2 zwei Funktionen $\varkappa(a_1, a_2)$ und $\lambda(a_1, a_2)$ einführen und beachten, daß

$$\frac{\partial f}{\partial \varkappa} = \frac{\partial f}{\partial a_1}\frac{\partial a_1}{\partial \varkappa} + \frac{\partial f}{\partial a_2}\frac{\partial a_2}{\partial \varkappa}, \qquad \frac{\partial f}{\partial \lambda} = \frac{\partial f}{\partial a_1}\frac{\partial a_1}{\partial \lambda} + \frac{\partial f}{\partial a_2}\frac{\partial a_2}{\partial \lambda}$$

ist. Es gibt nur eine Gleichung für die adiabatische Invariante (vgl. (1.9)):

$$\frac{\partial J}{\partial c}\cdot\overline{\frac{\partial f}{\partial \varkappa}} - \frac{\partial J}{\partial t_0}\cdot\overline{\frac{\partial f_2}{\partial \varkappa}} + \frac{\partial J}{\partial \varkappa} = 0, \tag{7.4}$$

indem die auf λ bezügliche Gleichung nach unserer Voraussetzung wegfällt. Wir wählen der Einfachheit halber $T = \infty$; dann sind $\varkappa, \lambda$ so zu wählen, daß mindestens einer der Mittelwerte in (7.4) existiert, d. h. entweder in $\frac{\partial f}{\partial \varkappa}$ oder $\frac{\partial f_2}{\partial \varkappa}$ die Glieder mit $t - t_0$ wegfallen, also

$$-a_2\frac{\partial a_1}{\partial \varkappa} + a_1\frac{\partial a_2}{\partial \varkappa} = 0 \quad \text{oder} \quad -a_1\frac{\partial a_1}{\partial \varkappa} + a_2\frac{\partial a_2}{\partial \varkappa} = 0 \tag{7.5}$$

wird. Setzen wir noch fest, daß etwa

$$\frac{\partial(\varkappa, \lambda)}{\partial(a_1, a_2)} = 1$$

sein soll, so gibt die erste Gleichung (7.5) die Werte

$$\varkappa = \frac{1}{2}a_1^2, \qquad \lambda = \frac{a_2}{a_1}, \tag{7.6}$$

und die zweite die Transformation

$$\varkappa = \frac{1}{4}\ln\frac{a_1 - a_2}{a_1 + a_2}, \qquad \lambda = a_1^2 - a_2^2 = \gamma^2. \tag{7.7}$$

Im Falle (7.6) erhalten wir aus (7.4) die Gleichung

$$-c \cdot \frac{\partial J}{\partial c} + 4\varkappa(1-\lambda^2) \cdot \frac{\partial J}{\partial \varkappa} = 0,$$

$$J = f(a_1 c^{2(1-\lambda^2)}) = f\left(a_1 c_1^2 c_1^{-2\left(\frac{a_2}{a_1}\right)^2}\right);$$

dies ist also die Invariante des gedämpften Oszillators, wenn das Verhältnis $\frac{a_2}{a_1}$ *absolut konstant bleibt.* Im Falle (7.7) führt die Gleichung

$$\mathfrak{Cof}\, 2\varkappa \cdot \frac{\partial J}{\partial t_0} - \sqrt{\lambda} \cdot \frac{\partial J}{\partial \varkappa} = 0$$

auf die Invariante

$$J = f\left(t_0 + \frac{1}{2\sqrt{\lambda}} \mathfrak{Sin}\, 2\varkappa\right) = f(2\gamma^2 t_0 - a_2),$$

die dann als solche gilt, wenn die reduzierte Frequenz γ konstant bleibt.

II.

§ 8.

Klassifizierung der n-dimensionalen Systeme.

Auf den im vorangehenden entwickelten Ausführungen aufbauend wenden wir uns nun der Behandlung n-dimensionaler Differentialsysteme vom Typus (1.1) zu und werden zunächst eine Klassifizierung der sich darbietenden Probleme vornehmen. Eine erste Unterteilung ergibt sich daraus, daß das Mittelungsintervall T entweder endlich oder unendlich groß ist, also man die Definition (1.7) bzw. (1.10) zugrunde legt. Wir verwenden nämlich in unsern Betrachtungen als wesentlichen Bestandteil die Integrale (1.3) des Systems, die aus der Auflösung der eindeutigen Beziehungen (1.2) entstanden sind; aber die Eindeutigkeit sowohl, wie die Auflösbarkeit sind Eigenschaften im kleinen, die jeweils nur für ein beschränktes Intervall der Variablen $t - t_0$, und damit der Trajektorie gelten. Die in (1.7) bzw. (1.10) zugrunde gelegten Definitionen der Mittelbildung verlangen aber mehr, sie fordern die Kenntnis der Integrale f_i im großen, nämlich im ganzen Intervall T oder ∞. Solange T endlich ist und in diesem Intervall die X_i regulär sind, ist diese Unterscheidung nicht wesentlich, denn ihm entspricht auf der Trajektorie $\mathfrak{T}$ ein endliches Stück S, das wir als singularitätenfrei voraussetzen wollen, und das durch eine endliche Anzahl der oben erwähnten Eindeutigkeitsbereiche überdeckt wird. Längs dieses eindimensionalen Bogens S sind die f_i also entweder eindeutig oder höchstens endlichvieldeutig und man kann S als endliches Stück der

Schnittlinie der $n-1$ Hyperflächen

$$f_i(x_1 \ldots x_n \mid a_1 \ldots a_\varrho) = c_i \qquad (i = 1 \ldots n-1) \tag{8.1}$$

des R_n deuten. Durch Projektion auf die Koordinatenebenen $(x_i\, x_k)$ geht S in ein endliches Bogenstück der $x_i\, x_k$-Ebene über; wir werden nun später gerade diese Projektionen zu betrachten haben und es ist daher naturgemäß, daß wir auf sie die im § 4 bezüglich des zweidimensionalen Falles gemachte Forderung übertragen, dergemäß die Endpunkte des Mittelungsintervalls von einer festen, die Parameter nicht enthaltenden Kurve ausgeschnitten werden. In den R_n zurückprojiziert besagt dies, daß die Endpunkte von S auf einer parameterfreien Hyperfläche:

$$\Phi(x_1 \ldots x_n) = 0$$

liegen müssen. Wir fassen also zusammen als ersten Fall:

A. *T ist endlich, S ebenfalls, die f_i längs S sind endlichvieldeutig. Die Endpunkte von S liegen auf einer parameterfreien Hyperfläche.*

Dieser Fall entspricht also im großen und ganzen den bisherigen Betrachtungen. Wir nehmen nun zweitens an, es sei $T = \infty$. Wie schon oben bemerkt, kommt es uns auf das Verhalten der f_i im großen an, und darüber ist bis heutigen Tags sehr wenig bekannt. Unsere Kenntnis darüber beschränkt sich im wesentlichen auf den Poincaré-Carathéodoryschen Wiederkehrsatz, der folgendes besagt: Liegen die Trajektorien von (1.1), wenigstens für die in Betracht kommenden Werte der Konstanten und Parameter, in einem zusammenhängenden Gebiet G des R_n von endlichem Maße, und kann ein Multiplikator μ gefunden werden, der in G positiv ist und höchstens in einer Nullmenge verschwindet, so kommen fast alle Trajektorien jedem ihrer Punkte unendlich oft beliebig nahe[12]). Es liegt daher nahe, anzunehmen, daß in diesem Falle die Trajektorien entweder geschlossen, also periodisch sind, oder fast alle Trajektorien eine gewisse Mannigfaltigkeit Φ dicht erfüllen. Die Integrale (1.3), die diese Trajektorien bestimmen, spalten sich dann, eventuell nach vorheriger Kombination, in zwei irreduzible Gruppen, deren erste diejenigen m unter ihnen umfaßt, die längs der ganzen Trajektorie $\mathfrak{T}$ eindeutig oder endlichvieldeutig sind:

$$f_\lambda(x_1 \ldots x_n \mid a_1 \ldots a_\varrho) = c_\lambda \qquad (\lambda = 1 \ldots m) \tag{8.2}$$

und deren zweite die längs $\mathfrak{T}$ unendlichvieldeutigen Integrale enthält, die also praktisch belanglos sind. Die Gleichungen (8.2) bestimmen im R_n

[12]) Vgl. Poincaré 1, Kap. 26; Carathéodory 1. Über das Wenige, das darüber hinaus bekannt ist, orientiert Smekal 1, S. 179–181; Levi-Civita 1, S. 331.

eine $n - m$-dimensionale Mannigfaltigkeit Φ, und wir werden uns nur mit solchen Systemen befassen, die die sogenannte *quasi-ergodische Hypothese* befriedigen: *Fast alle Trajektorien erfüllen die Mannigfaltigkeit Φ dicht.* Es gibt solche quasiergodische Systeme[13]). Bezüglich Φ wollen wir annehmen, daß es eine geschlossene Mannigfaltigkeit des R_n bilde.

Die Zahl m, die eine charakteristische Konstante des Systems bildet, bezeichnen wir als seine Imprimitivitätsordnung[14]). Die Trajektorien eines primitiven Systems erfüllen also ein Stück des R_n dicht und besitzen gar kein eindeutiges Integral, bei ihnen hat die Frage nach adiabatischen Invarianten offenbar keinen Sinn. Bei m-fach imprimitiven Systemen wird man nach solchen Invarianten suchen, die die m Konstanten $c_1 \ldots c_m$ mit den Parametern $a_1 \ldots a_\varrho$ in Beziehung setzen, und demnach, einer Bemerkung am Schluß des § 1 zufolge, in den Gleichungen (1.11) nur die diesbezüglichen Glieder berücksichtigen. Der größte Wert $m = n - 1$ wird angenommen, wenn alle Integrale des Systems endlichvieldeutig sind; Φ reduziert sich dann auf den eindimensionalen Schnitt der Hyperflächen (8.1), d. h. auf $\mathfrak{T}$, und somit ist in diesem Falle die Quasiergodenhypothese trivial. Wir fassen zusammen:

B. *T ist unendlich; das System ist von der Imprimitivitätsordnung m und erfüllt eine geschlossene $n - m$-dimensionale Mannigfaltigkeit quasiergodisch.*

Wir werden in drei Unterfälle teilen:

B 1. $m = 1$, das System ist einfach imprimitiv;

B 2. $1 < m < n - 1$, das System ist m-fach imprimitiv;

B 3. $m = n - 1$, alle Integrale f_i sind endlichvieldeutig, $\Phi \equiv \mathfrak{T}$ ist eindimensional.

Schließlich bietet sich als letzte Möglichkeit die dar, daß das System keine Quasiergodizitätseigenschaft besitzt, wie dies insbesondere bei Auftreten von Singularitäten zutreffen kann; solche Systeme entziehen sich, soweit man nicht auf Grund der Kenntnis ihrer Integrale ebenso verfahren kann, wie wir es in den Fällen A und B tun werden, unserer Behandlung; wir müssen sie im folgenden ausschließen.

C. *T ist unendlich; das System besitzt keine Quasiergodizitätseigenschaften.*

An Hand dieser Klassifikation wenden wir uns nun dem Problem der adiabatischen Invarianten zu.

---

[13]) Beispiele enthalten: Cherry 1; Levi-Civita 1, S. 326–328.

[14]) Wir schließen uns damit der Bezeichnung von Levi-Civita 1, S. 330 an, während sie in Levi-Civita 2 um eine Einheit verschoben ist.

§ 9.

Problemstellung für einfach-imprimitive Systeme.

Die im vorangehenden erörterten topologischen Voraussetzungen bezüglich der Trajektorien haben nur den einen Zweck, die in (1.10) definierte, auf t bezügliche Funktionaloperation in eine räumliche, auf Φ bezügliche Operation zu verwandeln, in ähnlicher Weise, wie wir es in (3.8) für zweidimensionale Systeme durchführen konnten. Die wirklich genuine Generalisierung der zweidimensionalen Systeme sind die einfach-imprimitiven; alle übrigen Fälle werden wir durch Reduktion mittels der bekannten Integrale auf zweidimensionale oder einfach-imprimitive Systeme zurückführen können.

Das System (1.1) sei einfach-imprimitiv und besitze das endlichvieldeutige Integral

$$f(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) = c, \tag{9.1}$$

durch das die $n-1$-dimensionale geschlossene Hyperfläche Φ charakterisiert wird. Die Verteilungsfunktion F wollen wir für den Augenblick als konstant annehmen. Wir wollen dann erreichen, daß an Stelle der Definition (1.10) eine Mittelbildung, d. h. eine Integration, über Φ tritt, was zufolge der quasiergodischen Verteilung der Trajektorien über Φ möglich ist. Es fragt sich nur, welche Verteilungsfunktion müssen wir dieser Mittelbildung zuordnen oder welches ist die eindeutige Dichte $\varkappa$, mit der wir jedes Flächenelement $d\Phi$ zu multiplizieren haben?

$\varkappa$ muß eine gewisse Invarianzeigenschaft erfüllen, es muß nämlich „auf den Trajektorien mitgenommen werden", d. h. sind $d\Phi$, $d\Phi_1$ zwei Elemente von Φ, die infolge der Gleichungen (1.1) auseinander hervorgehen (in dem Sinne, daß, wenn allen Punkten von $d\Phi$ der Wert $t = t_0$ zugeordnet wird, $d\Phi_1$ die Gesamtheit der Punkte darstellt, die auf den entsprechenden Trajektorien zum Werte $t = t_1$ gehören), so muß

$$\varkappa\, d\Phi = \varkappa_1\, d\Phi_1 \tag{9.2}$$

sein; dies folgt daraus, daß der Mittelwert (1.10) unabhängig von t_0, d. h. vom Ausgangselement $d\Phi$ sein muß.

Wir zeigen nun, daß es im wesentlichen *nur eine* Funktion $\varkappa$ gibt, die dieser Invarianz genügt. Gälte nämlich dasselbe von einer andern eindeutigen Dichteverteilung λ, so müßten $\varkappa\, d\Phi$, $\lambda\, d\Phi$ also auch $\varkappa : \lambda$ gegenüber (1.1) invariant, mithin letztere Größe ein Integral des Systems (1.1) sein. Da dieses aber nach Voraussetzung nur ein endlichvieldeutiges Integral, nämlich f, besitzt, muß also

$$\varkappa : \lambda = \Psi(f(x_1 \dots x_n \,|\, a_1 \dots a_\varrho)),$$

also dieses Verhältnis auf Φ konstant sein; auf einen Proportionalitätsfaktor kommt es aber bei der Dichte nicht an, demnach gibt es im wesentlichen nur eine Dichtefunktion $\varkappa$.[15])

Man kann die $(n-1)$-dimensionalen Flächenelemente $d\Phi$, $d\Phi_1$ leicht zu Raumelementen $d\tau$, $d\tau_1$ des R_n ergänzen, die auseinander infolge von (1.1) hervorgehen. Dazu betrachten wir neben der Fläche (9.1) die Hyperfläche Φ' mit der Gleichung

$$f(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) = c + dc; \tag{9.3}$$

wir ordnen jedem Punkte P von Φ den auf seiner Normalen liegenden Punkt P' von Φ' zu und bezeichnen den Vektor PP' mit $d'n$; diese Zuordnung ist, weil wir Φ als geschlossen und frei von Singularitäten annehmen, überall eindeutig. Die Richtungkosinus der Normalen sind

$$\frac{1}{\Gamma_{1n}} \frac{\partial f}{\partial x_i} \quad (i = 1 \dots n); \qquad \Gamma_{1n}^2 = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2, \tag{9.4}$$

und daher die Komponenten von $d'n$:

$$d'x_i = \frac{1}{\Gamma_{1n}} \frac{\partial f}{\partial x_i} d'n;$$

andrerseits folgt aus (9.1), (9.3) und (9.4):

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d'x_i = \Gamma_{1n} d'n = dc,$$

mithin

$$d'n = \frac{dc}{\Gamma_{1n}}; \qquad d'x_i = \frac{1}{\Gamma_{1n}^2} \frac{\partial f}{\partial x_i} dc. \tag{9.5}$$

Nennen wir nun $d\tau = d\Phi\, d'n$ das über $d\Phi$ liegende Raumelement der zwischen Φ und Φ' liegenden Schicht, so geht es infolge der Gleichungen (1.1) in das über $d\Phi_1$ liegende Schichtelement $d\tau_1 = d\Phi_1 d'n_1$ über, da die in der Schicht beginnenden Trajektorien dieselbe nie verlassen können und überall dicht ausfüllen. Aus (9.2) folgt dann, daß

$$\varkappa\, d\Phi\, dc = \varkappa \Gamma_{1n}\, d\tau = \varkappa_1 d\Phi_1 dc = (\varkappa \Gamma_{1n})_1 d\tau_1,$$

also die Größe

$$\mu = \varkappa \Gamma_{1n} \tag{9.6}$$

die Beziehung befriedigen muß:

$$\mu\, d\tau = \mu_1 d\tau_1,$$

d. h. in moderner Ausdrucksweise: $\int \mu\, d\tau$ *ist eine Integralinvariante des Systems* (1.1). Nun gilt der leicht zu beweisende Satz[16]), daß, wenn diese

[15]) Vgl. Levi-Civita 1, S. 337f.

[16]) Vgl. Poincaré 1, S. 41ff.

Größe eine Integralinvariante ist, dann der Faktor μ ein Jacobischer Multiplikator von (1.1) sein, d. h. die Differentialgleichung

$$\sum_{i=1}^{n} \frac{\partial \mu}{\partial x_i} X_i + \mu \sum_{i=1}^{n} \frac{\partial X_i}{\partial x_i} = 0 \tag{9.7}$$

erfüllen muß, und umgekehrt gibt jeder Multiplikator eine Integralinvariante. Mithin folgt aus (9.6)

$$\varkappa = \frac{\mu}{\Gamma_{1n}}, \tag{9.8}$$

wo μ jeden beliebigen eindeutigen bzw. endlichvieldeutigen Multiplikator von (1.1) bedeuten kann. Es gibt im wesentlichen nur einen solchen Multiplikator μ, da der Quotient zweier solcher ein endlichvieldeutiges Integral von (1.1), also nach Voraussetzung eine bloße Funktion von f und demnach auf Φ konstant ist.

Wir setzen im folgenden voraus, daß es eine endlichvieldeutige Lösung μ von (9.7) gibt und daß μ auf Φ — abgesehen von endlich vielen Punkten — nicht verschwindet, also ein bestimmtes Vorzeichen hat, das wir als positiv wählen. Wir werden später (§ 11) sehen, daß diese Voraussetzung gleichbedeutend ist mit der Aussage: Die die Trajektorie im kleinen definierenden Funktionen $f_1 \dots f_{n-1}$ sollen längs derselben keine Singularitäten besitzen. An die Stelle der Mittelbildung (1.10) tritt dann also (immer noch $F \equiv 1$ angenommen) die Definition:

$$\bar{\alpha} = \int_{\Phi} \alpha(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) \frac{\mu}{\Gamma_{1n}} d\Phi : \int_{\Phi} \frac{\mu}{\Gamma_{1n}} d\Phi. \tag{9.9}$$

Ist $F \not\equiv 1$, so tritt hier noch unter das Integralzeichen die durch (1.3) transformierte, endlichvieldeutige Verteilungsfunktion $F(x_1 \dots x_n \,|\, a_1 \dots a_\varrho)$ hinzu, so daß

$$\bar{\alpha} = \int_{\Phi} \alpha(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) F(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) \frac{\mu}{\Gamma_{1n}} d\Phi : \int_{\Phi} F(x_1 \dots x_n \,|\, a_1 \dots a_\varrho) \frac{\mu}{\Gamma_{1n}} d\Phi \tag{9.10}$$

wird.

Damit gehen wir in die Differentialgleichungen (1.11) der adiabatischen Invarianten ein und berücksichtigen nur die auf c und a_ν bezüglichen Differentiationen. Dann lautet dieses Differentialsystem:

$$\frac{\partial J}{\partial c} \cdot \int_{\Phi} \frac{\partial f}{\partial a_\nu} F \frac{\mu}{\Gamma_{1n}} d\Phi + \frac{\partial J}{\partial a_\nu} \cdot \int_{\Phi} F \frac{\mu}{\Gamma_{1n}} d\Phi = 0 \qquad (\nu = 1 \dots \varrho), \tag{9.11}$$

das die evidente Generalisierung von (3.10) bildet. Wir werden seine Integrabilitätsbedingungen und seine Lösung später untersuchen.

§ 10.

Problemstellung für m-fach imprimitive Systeme.

Ist das System (1.1) m-fach imprimitiv, d.h. gestattet es die m endlichvieldeutigen Integrale

$$f_\lambda(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho) = c_\lambda \qquad (\lambda = 1 \ldots m), \tag{10.1}$$

die die geschlossene $n-m$-dimensionale Mannigfaltigkeit Φ bestimmen, so führen ähnliche Überlegungen wie oben zum Ziel. Die Dichtefunktion $\varkappa$ muß wieder der Invarianzeigenschaft (9.2) genügen und ist damit bis auf einen konstanten Faktor auf Φ bestimmt, da der Quotient zweier solcher Dichten ein endlichvieldeutiges Integral des Systems, also eine Funktion der $f_1 \ldots f_m$ sein muß.

Wir ergänzen wieder $d\Phi$, $d\Phi_1$ zu Raumelementen des R_n und verfahren dazu folgendermaßen: Neben Φ führen wir die Mannigfaltigkeit Φ' ein, die durch die Gleichungen

$$f_\lambda(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho) = c_\lambda + dc_\lambda \qquad (\lambda = 1 \ldots m) \tag{10.2}$$

definiert wird und mit Φ zusammen eine von den Trajektorien dicht ausgefüllte Schicht bildet, die von diesen nicht verlassen wird. Die Normalen auf den Hyperflächen (10.1) haben die Richtungskosinus

$$\frac{\partial f_\lambda}{\partial x_i} : \sqrt{\sum_{i=1}^{n} \left(\frac{\partial f_\lambda}{\partial x_i}\right)^2} \qquad (i = 1 \ldots n;\ \lambda = 1 \ldots m), \tag{10.3}$$

sie stehen natürlich auf jedem in Φ enthaltenen Vektor senkrecht. Wir bestimmen ferner $n-m$ weitere Funktionen $u_j(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho)$ derart, daß die Hyperflächen

$$u_j(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho) = \mathfrak{c}_j \qquad (j = 1 \ldots n-m) \tag{10.4}$$

sämtliche Flächen (10.1) orthogonal schneiden und selbst zueinander orthogonal sind, d. h. daß

$$\sum_{i=1}^{n} \frac{\partial u_j}{\partial x_i} \frac{\partial f_\lambda}{\partial x_i} = 0 \qquad (j = 1 \ldots n-m,\ \lambda = 1 \ldots m), \tag{10.5}$$

$$\sum_{i=1}^{n} \frac{\partial u_j}{\partial x_i} \frac{\partial u_k}{\partial x_i} = 0 \qquad (j \neq \varkappa = 1 \ldots n-m) \tag{10.6}$$

gilt. Eine solche Wahl der Funktionen u_j ist stets möglich. Die Normalenrichtungen auf (10.4) berühren zufolge (10.5) die Mannigfaltigkeit Φ und bilden demnach in einem Punkte deren $n-m$-dimensionale Tangentialmannigfaltigkeit. Betrachten wir neben den Flächen (10.4) noch die folgenden:

$$u_j(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho) = \mathfrak{c}_j + d\mathfrak{c}_j \qquad (j = 1 \ldots n-m), \tag{10.7}$$

so gilt für je ein Paar zusammengehöriger Flächen (10.4) und (10.7) nach (9.5), daß das dazwischen liegende Normalenstück die Länge

$$d' n_j = \left\{ \sum_{i=1}^{n} \left(\frac{\partial u_j}{\partial x_i} \right)^2 \right\}^{-\frac{1}{2}} d\mathfrak{c}_j$$

besitzt. Das Element $d\Phi$ ist definitionsgemäß gleich dem seiner Tangentialmannigfaltigkeit, und dessen Größe wiederum ist zufolge der Orthogonalität der $d' n_i$ gleich

(10.8) $$d\Phi = \prod_{j=1}^{n-m} d' n_j = \left\{ \prod_{j=1}^{n-m} \sum_{i=1}^{n} \left(\frac{\partial u_j}{\partial x_i} \right)^2 \right\}^{-\frac{1}{2}} d\mathfrak{c}_1 \ldots d\mathfrak{c}_{n-m}.$$

Die $2n$ Flächen (10.1), (10.2), (10.4) und (10.7) definieren in ihrem Schnitt ein Raumelement $d\tau$ des R_n; die lineare Unabhängigkeit der Funktionen f_λ, u_j gestattet, diese Größen als neue Koordinaten des R_n einzuführen und demnach zu setzen:

(10.9) $$d\tau = dx_1 \ldots dx_n = \left[\frac{\partial (f_1 \ldots f_m, u_1 \ldots u_{n-m})}{\partial (x_1 \ldots x_n)} \right]^{-1} . dc_1 \ldots dc_m \, d\mathfrak{c}_1 \ldots d\mathfrak{c}_{n-m}.$$

Andrerseits folgert man aus (10.5), (10.6), daß:

$$\left[\frac{\partial (f_1 \ldots f_m, u_1 \ldots u_{n-m})}{\partial (x_1 \ldots x_n)} \right]^2 = \prod_{j=1}^{n-m} \sum_{i=1}^{n} \left(\frac{\partial u_j}{\partial x_i} \right)^2 \cdot \Gamma^2_{m,n},$$

worin

$$\Gamma^2_{m,n} = \operatorname*{Det}_{\varkappa,\lambda} \sum_{i=1}^{n} \frac{\partial f_\varkappa}{\partial x_i} \frac{\partial f_\lambda}{\partial x_i} = \left(\frac{\partial (f_1 \ldots f_m)}{\partial (x_1 \ldots x_n)} \right)^2$$

das Quadrat der letzthingeschriebenen Matrix bedeutet. (10.9) gibt also in Verbindung mit (10.8)

(10.10) $$d\tau = \frac{1}{\Gamma_{m,n}} d\Phi \cdot dc_1 \ldots dc_m$$

als Ausdruck für das Raumelement der zwischen Φ und Φ' liegenden Schicht.

Aus (9.2) ergibt sich also

$$\varkappa \, d\Phi \, dc_1 \ldots dc_m = \varkappa \, \Gamma_{m,n} \, d\tau = \varkappa_1 \, d\Phi_1 \, dc_1 \ldots dc_m = (\varkappa \, \Gamma_{m,n})_1 \, d\tau_1,$$

woraus man, wie in § 9, folgert, daß die Größe

(10.11) $$\mu = \Gamma_{m,n} \varkappa$$

ein Multiplikator des vorgelegten Systems sein muß, und umgekehrt gibt jeder auf Φ endlichvieldeutige Multiplikator μ — der durch diese Forderung bis auf einen konstanten Faktor festgelegt ist — durch (10.11) eine gültige Dichtefunktion

$$\varkappa = \frac{\mu}{\Gamma_{m,n}}.$$

Die an Stelle von (1.10) anzuwendende Funktionaloperation lautet sonach in diesem Falle:

(10.12) $$\bar{\alpha} = \int\limits_{\Phi} \alpha F \frac{\mu}{\Gamma_{m,n}} d\Phi : \int\limits_{\Phi} F \frac{\mu}{\Gamma_{m,n}} d\Phi,$$

und die später zu behandelnden Differentialgleichungen der adiabatischen Invarianten werden:

(10.13) $$\sum_{\lambda=1}^{m} \frac{\partial J}{\partial c_\lambda} \cdot \int\limits_{\Phi} \frac{\partial f_\lambda}{\partial a_\nu} F \frac{\mu}{\Gamma_{m,n}} d\Phi + \frac{\partial J}{\partial a_\nu} \cdot \int\limits_{\Phi} F \frac{\mu}{\Gamma_{m,n}} d\Phi = 0 \qquad (\nu = 1 \ldots \varrho).$$

§ 11.

$n-1$-fach imprimitive Systeme.

Der Klassifikation des § 8 entsprechend bleiben noch die Fälle A und B 3 zu besprechen, die unabhängig von allen topologischen Voraussetzungen und der Quasiergodenhypothese sind. Man kennt die Integrale

(11.1) $$f_i(x_1 \ldots x_n \mid a_1 \ldots a_\varrho) = c_i \qquad (i = 1 \ldots n-1)$$

und folgert aus den Identitäten

(11.2) $$\sum_{\varkappa=1}^{n} \frac{\partial f_i}{\partial x_\varkappa} X_\varkappa = 0 \qquad (i = 1 \ldots n-1)$$

die Relationen:

(11.3) $$X_i = \frac{(-1)^i}{\mu} \frac{\partial(f_1 \ldots f_{n-1})}{\partial(x_1 \ldots x_{i-1} x_{i+1} \ldots x_n)} = \frac{(-1)^i}{\mu} \mathfrak{F}_i,$$

worin μ einen Proportionalitätsfaktor bedeutet und wir gesetzt haben:

$$\mathfrak{F}_i = \frac{\partial(f_1 \ldots f_{n-1})}{\partial(x_1 \ldots x_{i-1} x_{i+1} \ldots x_n)}.$$

Da die $\mathfrak{F}_i$ nicht unabhängig voneinander sind, muß μ eine leicht zu findende Gleichung befriedigen; es ist nämlich:

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i}(\mu X_i) = \sum_{i=1}^{n} (-1)^i \frac{\partial \mathfrak{F}_i}{\partial x_i}$$

$$= \sum_{i=1}^{n} \sum_{\varkappa=1}^{n-1} \sum_{l=1}^{i-1} (-1)^{\varkappa+l+i} \frac{\partial^2 f_\varkappa}{\partial x_l \partial x_i} \frac{\partial(f_1 \ldots f_{\varkappa-1} f_{\varkappa+1} \ldots f_{n-1})}{\partial(x_1 \ldots x_{l-1} x_{l+1} \ldots x_{i-1} x_{i+1} \ldots x_n)}$$

$$- \sum_{i=1}^{n} \sum_{\varkappa=1}^{n-1} \sum_{l=i+1}^{n} (-1)^{\varkappa+l+i} \frac{\partial^2 f_\varkappa}{\partial x_l \partial x_i} \frac{\partial(f_1 \ldots f_{\varkappa-1} f_{\varkappa+1} \ldots f_{n-1})}{\partial(x_1 \ldots x_{i-1} x_{i+1} \ldots x_{l-1} x_{l+1} \ldots x_n)}$$

$$\equiv 0,$$

wie man durch Vertauschung von i und l in der zweiten Summe erkennt; μ genügt also der Differentialgleichung

$$(11.4)\qquad \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(\mu X_i) = \sum_{i=1}^{n} \frac{\partial \mu}{\partial x_i} X_i + \mu \sum_{i=1}^{n} \frac{\partial X_i}{\partial x_i} = 0,$$

ist mithin ein Jacobischer Multiplikator des Systems (1.1). Führen wir dessen Divergenz

$$(11.5)\qquad \Delta_0 = \sum_{i=1}^{n} \frac{\partial X_i}{\partial x_i}$$

ein, so kann man (11.4) auch in die Form setzen

$$(11.6)\qquad \sum_{i=1}^{n} X_i \frac{\partial \log \mu}{\partial x_i} + \Delta_0 = 0.$$

Wir wollen auch hier, wie in den vorangehenden Paragraphen, die Forderung stellen, daß μ auf der Trajektorie (außer eventuell in endlich vielen Punkten) nicht verschwindet. Wäre μ längs $\mathfrak{T}$ Null, so besagte dies nach (11.3), da wir die X_i daselbst als regulär voraussetzen, daß längs $\mathfrak{T}$ sämtliche $\mathfrak{F}_i$ verschwinden, und dies wiederum, daß die Gleichungen (11.2) linear abhängig sind, also etwa eine Relation

$$\frac{\partial f_{n-1}}{\partial x_i} = \sum_{\varkappa=1}^{n-2} \alpha_\varkappa \frac{\partial f_\varkappa}{\partial x_i} \qquad (i = 1 \ldots n)$$

mit konstanten $\alpha_\varkappa$ besteht. Geometrisch drückt dies aus, daß die Integralfläche

$$f_{n-1} - \sum_{\varkappa=1}^{n-2} \alpha_\varkappa f_\varkappa = \text{konst.}$$

des Systems (1.1) längs $\mathfrak{T}$ singulär ist. Unsere Annahme fordert also, *daß keine mögliche Integralfläche die Trajektorie als singuläre Linie besitzt.* Die gleiche Überlegung hat auch für die vorangehenden Paragraphen Geltung, nur daß sie dort *im kleinen* gilt.

Nunmehr folgert man aus (1.1) und (11.3), daß

$$(11.7)\qquad dt = \frac{dx_i}{X_i} = (-1)^i \mu \frac{dx_i}{\mathfrak{F}_i} = \mu \sqrt{\frac{\sum_{i=1}^{n} dx_i^2}{\sum_{i=1}^{n} \mathfrak{F}_i^2}} = \mu \frac{ds}{\Gamma_{n-1,n}},$$

worin ds das Bogenelement der Trajektorie $\mathfrak{T}$ in R_n bezeichnet und gesetzt ist:

$$(11.8)\qquad \Gamma_{n-1,n}^2 = \sum_{i=1}^{n} \mathfrak{F}_i^2 = \sum_{i=1}^{n} \left[\frac{\partial(f_1 \ldots f_{n-1})}{\partial(x_1 \ldots x_{i-1}\, x_{i+1} \ldots x_n)}\right]^2.$$

Einer bekannten Identität zufolge stimmt $\Gamma_{n-1,n}$ mit dem Ausdruck überein, der aus der im vorigen Paragraphen definierten Größe $\Gamma_{m,n}$ für $m = n - 1$ hervorgeht; man kann ihr eine geometrische Bedeutung unterschieben durch die Bemerkung, daß

$$(-1)^i \frac{\mathfrak{F}_i}{\Gamma_{n-1,n}} = \mu \frac{X_i}{\Gamma_{n-1,n}}$$

die Richtungskosinus der Tangente an $\mathfrak{T}$ sind. Auf Grund von (11.7) tritt an Stelle von (1.7) die Mittelbildung:

$$\bar{\alpha} = \int\limits_S \alpha F \frac{\mu}{\Gamma_{n-1,n}} ds : \int\limits_S F \frac{\mu}{\Gamma_{n-1,n}} ds, \tag{11.9}$$

worin S das dem Intervall T von t entsprechende Stück der Trajektorie bzw. für $T = \infty$ die Gesamtlänge der letzteren bezeichnet. Die Differentialgleichungen der adiabatischen Invarianten lauten demnach in diesem Falle

$$\sum_{i=1}^{n-1} \frac{\partial J}{\partial c_i} \cdot \int\limits_S \frac{\partial f_i}{\partial a_\nu} F \frac{\mu}{\Gamma_{n-1,n}} ds + \frac{\partial J}{\partial a_\nu} \cdot \int\limits_S F \frac{\mu}{\Gamma_{n-1,n}} ds = 0 \quad (\nu = 1 \ldots \varrho). \tag{11.10}$$

§ 12.

Die Invarianten einfach-imprimitiver Systeme.

Nachdem so die Differentialgleichungen der adiabatischen Invarianten in allen Fällen aufgestellt sind, wollen wir zunächst ihre Integrabilitätsbedingungen und mögliche Lösung für einfach-imprimitive Systeme untersuchen. Bezeichnen wir analog zu (4.2):

$$g = \int\limits_\Phi F \frac{\mu}{\Gamma_{1n}} d\Phi, \qquad g_\nu = \int\limits_\Phi \frac{\partial f}{\partial a_\nu} F \frac{\mu}{\Gamma_{1n}} d\Phi, \tag{12.1}$$

so lauten die Integrabilitätsbedingungen des Systems (9.11) genau wie (4.3). Um sie in expliziter Form zu gewinnen, führen uns die §§ 4, 6 zu einem Kunstgriff.

Es sei im Innern von Φ eine stetige Ortsfunktion $G(x_1 \ldots x_n \mid a_1 \ldots a_\varrho)$ vorgelegt; wir betrachten dann das Integral

$$H = \int G(x_1 \ldots x_n \mid a_1 \ldots a_\varrho)\, dx_1 \ldots dx_n, \tag{12.2}$$

wo sich die Integration über das ganze Innere von Φ erstrecken soll, und wollen seine Ableitungen berechnen. Um $\frac{\partial H}{\partial c}$ zu finden, müssen wir neben Φ die durch (9.3) definierte Hyperfläche Φ' betrachten und haben mit den dortigen Bezeichnungen:

$$\frac{\partial H}{\partial c} = \int\limits_\Phi G\, d\Phi\, d'n : \dot{dc} = \int\limits_\Phi G \frac{1}{\Gamma_{1n}} d\Phi, \tag{12.3}$$

wie aus (9.5) folgt. Um dagegen $\frac{\partial H}{\partial a_\nu}$ zu ermitteln, führen wir neben Φ die Fläche Φ^* ein, deren Gleichung

(12.4) $$f(x_1 \ldots x_n \mid a_1 \ldots a_\nu + da_\nu \ldots a_\varrho) = c$$

lautet, und deren Punkte P^* wir uns aus denen P von Φ durch die normale Verschiebung d^*n entstanden denken, deren Richtungskosinus durch (9.4) gegeben sind, und deren Komponenten demnach lauten

$$d^* x_i = \frac{1}{\Gamma_{1n}} \frac{\partial f}{\partial x_i} d^* n;$$

andrerseits gibt uns (12.4):

$$\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} d^* x_i = \Gamma_{1n} d^* n = - \frac{\partial f}{\partial a_\nu} da_\nu,$$

also

(12.5) $$d^* n = - \frac{\partial f}{\partial a_\nu} \frac{1}{\Gamma_{1n}} da_\nu; \qquad d^* x_i = - \frac{1}{\Gamma_{1n}^2} \frac{\partial f}{\partial a_\nu} \frac{\partial f}{\partial x_i} da_\nu.$$

Bei der Berechnung von $\frac{\partial H}{\partial a_\nu}$ ist zu beachten, daß durch die Variation von a_ν nicht nur Φ, sondern auch die Funktion G im Innern von Φ variiert, so daß also

$$\frac{\partial H}{\partial a_\nu} = \int_\Phi G \, d\Phi \, d^* n : da_\nu + \int \frac{\partial G}{\partial a_\nu} dx_1 \ldots dx_n$$

(12.6) $$= - \int_\Phi \frac{\partial f}{\partial a_\nu} G \frac{1}{\Gamma_{1n}} d\Phi + \int \frac{\partial G}{\partial a_\nu} dx_1 \ldots dx_n$$

wird. Vergleicht man diese Formeln mit den Integralen g, g_ν, so liegt es also nahe, das Integral

(12.7) $$V = \int F \mu \, dx_1 \ldots dx_n,$$

über das Innere von Φ erstreckt, einzuführen. Dann ist nämlich nach (12.3) und (12.6)

(12.8) $$g = \frac{\partial V}{\partial c}, \qquad g_\nu = - \frac{\partial V}{\partial a_\nu} - \int \frac{\partial (F\mu)}{\partial a_\nu} dx_1 \ldots dx_n,$$

woraus folgt, indem man wieder (12.3) und (12.6) heranzieht, daß:

(12.9) $$\begin{aligned} \frac{\partial g_\varkappa}{\partial a_\lambda} - \frac{\partial g_\lambda}{\partial a_\varkappa} &= \int_\Phi \left\{ \frac{\partial f}{\partial a_\lambda} \frac{\partial (F\mu)}{\partial a_\varkappa} - \frac{\partial f}{\partial a_\varkappa} \frac{\partial (F\mu)}{\partial a_\lambda} \right\} \frac{1}{\Gamma_{1n}} d\Phi, \\ \frac{\partial g}{\partial a_\varkappa} + \frac{\partial g_\varkappa}{\partial c} &= - \int \frac{\partial (F\mu)}{\partial a_\nu} \frac{1}{\Gamma_{1n}} d\Phi \end{aligned}$$

gilt.

Die Integrabilitätsbedingungen (4.3) lauten demzufolge explizit:

$$
\begin{aligned}
&\int_\Phi F\frac{\mu}{\Gamma_{1n}}d\Phi\cdot\int_\Phi\left\{\frac{\partial f}{\partial a_\lambda}\frac{\partial(F\mu)}{\partial a_\varkappa}-\frac{\partial f}{\partial a_\varkappa}\frac{\partial(F\mu)}{\partial a_\lambda}\right\}\frac{1}{\Gamma_{1n}}d\Phi\\
(12.10)\qquad &+\int_\Phi\frac{\partial f}{\partial a_\varkappa}F\frac{\mu}{\Gamma_{1n}}d\Phi\cdot\int_\Phi\frac{\partial(F\mu)}{\partial a_\lambda}\frac{1}{\Gamma_{1n}}d\Phi\\
&-\int_\Phi\frac{\partial f}{\partial a_\lambda}F\frac{\mu}{\Gamma_{1n}}d\Phi\cdot\int_\Phi\frac{\partial(F\mu)}{\partial a_\varkappa}\frac{1}{\Gamma_{1n}}d\Phi=0\qquad(\varkappa,\lambda=1\ldots\varrho).
\end{aligned}
$$

Unter Benutzung der Bezeichnung (9.10) können wir somit den Satz formulieren:

Satz 7. *Die notwendigen und hinreichenden Bedingungen für die Existenz einer adiabatischen Invariante eines einfach-imprimitiven Systems lauten*:

$$
(12.11)\qquad \overline{\frac{\partial f}{\partial a_\varkappa}\frac{\partial\log\mu F}{\partial a_\lambda}}-\overline{\frac{\partial f}{\partial a_\varkappa}}\cdot\overline{\frac{\partial\log\mu F}{\partial a_\lambda}}=\overline{\frac{\partial f}{\partial a_\lambda}\frac{\partial\log\mu F}{\partial a_\varkappa}}-\overline{\frac{\partial f}{\partial a_\lambda}}\cdot\overline{\frac{\partial\log\mu F}{\partial a_\varkappa}}
$$
$$(\varkappa,\lambda=1\ldots\varrho).$$

Ein Vergleich mit (4.30) zeigt, daß sie die gleiche Form haben wie im zweidimensionalen Falle, und daher lassen sich entsprechende Erörterungen anknüpfen wie dort.

Die Voraussetzung, daß Φ im R_n geschlossen sei, ermöglicht die Feststellung, daß auf Φ sowohl die Koordinaten $x_1\ldots x_{n-1}$, als auch alle Funktionen derselben, im besonderen $\frac{\partial f}{\partial a_\varkappa}$ und $\frac{\partial\log\mu F}{\partial a_\varkappa}$, nach $n-1$ reellen Parametern $t_1\ldots t_{n-1}$ in mehrfache Fourierreihen entwickelbar sind. Setzen wir dann

$$
(12.12)\qquad \frac{\partial f}{\partial a_\varkappa}=\alpha_\varkappa+h_\varkappa(x_1\ldots x_n),\qquad \frac{\partial\log\mu F}{\partial a_\varkappa}=\beta_\varkappa+j_\varkappa(x_1\ldots x_n)
$$
$$(\varkappa=1\ldots\varrho),$$

worin

$$\alpha_\varkappa=\overline{\frac{\partial f}{\partial a_\varkappa}},\qquad \beta_\varkappa=\overline{\frac{\partial\log\mu F}{\partial a_\varkappa}}$$

die konstanten Terme der Fourierreihen für $F\frac{\partial f}{\partial a_\varkappa}$ und $F\frac{\partial\log\mu F}{\partial a_\varkappa}$ bedeuten, so sind die $h_\varkappa$, $j_\varkappa$ Funktionen, deren Mittelwert über Φ verschwindet, und man kann (12.11) die Form geben:

$$(12.13)\qquad \overline{h_\varkappa j_\lambda}=\overline{h_\lambda j_\varkappa},$$

was sich auch als Relation zwischen Fourierkonstanten schreiben läßt.

Drei Spezialfälle treten denen des § 5 an die Seite; die Entwicklungen sind die gleichen wie dort, wir sprechen nur die Resultate aus:

Fall I. $$h_\varkappa(x_1 \ldots x_n) = 0 \qquad (\varkappa = 1 \ldots \varrho).$$

Aus den Gleichungen (1.8) verschwindet die Funktionaloperation, sie werden mit (1.6) identisch. f ist ein gegenüber den Parametern statio näres Integral, d. h. nicht nur Integral des Systems (1.1), sondern auch der ϱ Systeme

$$\frac{dx_i}{dt} = \frac{\partial X_i}{\partial a_\varkappa} \qquad (\varkappa = 1 \ldots \varrho), \tag{12.14}$$

wofür notwendig ist, daß in der Matrix

$$\begin{pmatrix} X_1 & \ldots & X_n \\ \frac{\partial X_1}{\partial a_\varkappa} & \ldots & \frac{\partial X_n}{\partial a_\varkappa} \end{pmatrix} \qquad (\varkappa = 1 \ldots \varrho)$$

alle n-reihigen Determinanten verschwinden.

Fall II. Für einen bestimmten Multiplikator μ ist

$$j_\varkappa(x_1 \ldots x_n) = 0 \qquad (\varkappa = 1 \ldots \varrho),$$

also

$$\frac{\partial \log \mu F}{\partial a_\varkappa} = \Psi_\varkappa(f | a_1 \ldots a_\varrho) \qquad (\varkappa = 1 \ldots \varrho). \tag{12.15}$$

Für $F \equiv 1$ folgt, daß μ ein stationärer Multiplikator, d. h. auch ein solcher der Systeme (12.14) ist. Notwendig ist dazu, daß in der Matrix

$$\begin{pmatrix} X_1 & \ldots & X_n, & \Delta_0 \\ \frac{\partial X_1}{\partial a_\varkappa} & \ldots & \frac{\partial X_n}{\partial a_\varkappa}, & \frac{\partial \Delta_0}{\partial a_\varkappa} \end{pmatrix} \qquad (\varkappa = 1 \ldots \varrho)$$

sämtliche $(n+1)$-reihigen Determinanten verschwinden. Dies ist im besonderen für Liouvillesche Systeme ($\Delta_0 = 0$), speziell Hamiltonsche Systeme, der Fall.

Fall III. $$j_\varkappa(x_1 \ldots x_n) = \text{konst.}\, h_\varkappa(x_1 \ldots x_n) \qquad (\varkappa = 1 \ldots \varrho)$$

läßt sich wie in § 5 auf den vorangehenden Fall zurückführen.

Zusammenfassend tritt also dem Satz 4 der folgende an die Seite:

Satz 8. *Hinreichende Bedingungen für die Existenz der adiabatischen Invariante eines einfach-imprimitiven Systems sind*: a) *die Stationarität des endlichvieldeutigen Integrals*, b) *bei einfacher Mittelbildung die Stationarität eines Multiplikators gegenüber den Parametern.*

§ 13.

Die generalisiert-Liouvilleschen Systeme.

In (12.15) ist als besondere Möglichkeit die enthalten, daß

$$\frac{\partial \mu F}{\partial a_\varkappa} = 0 \qquad (\varkappa = 1 \ldots \varrho) \tag{13.1}$$

ist; wir wollen dann das System als generalisiert-Liouvillesches bezeichnen.

Aus (12.8) einerseits und (9.11) andrerseits folgt, daß die Gleichungen der adiabatischen Invariante die Form annehmen

$$\frac{\partial J}{\partial c}\frac{\partial V}{\partial a_\nu} - \frac{\partial J}{\partial a_\nu}\frac{\partial V}{\partial c} = 0 \qquad (\nu = 1 \ldots \varrho),$$

d. h. aber:

$$J = V = \int F\mu\, dx_1 \ldots dx_n \tag{13.2}$$

ist die adiabatische Invariante unseres Systems. Dem Satz 5 tritt also der folgende an die Seite:

Satz 9. *Die adiabatische Invariante eines einfach-imprimitiven generalisiert-Liouvilleschen Systems ist*

$$V = \int F\mu\, dx_1 \ldots dx_n.$$

Handelt es sich um einfache Mittelbildung ($F = 1$), so ist in der betrachteten Systemklasse speziell die Klasse der Liouvilleschen Systeme enthalten, für die $\Delta_0 = 0$, $\mu =$ konst. ist; für sie reduziert sich die Invariante auf das von Φ eingeschlossene Volumen des R_n. Als besondere Möglichkeit sind hierin wieder diejenigen Hamiltonschen Systeme enthalten, die außer dem Energieintegral kein weiteres endlichvieldeutiges Integral besitzen und für die unser Satz mit der Aussage des Gibbs-Hertzschen Theorems[17]) zusammenfällt, daß das Phasenvolumen der Energiefläche adiabatisch invariant ist. Allgemeiner umfaßt (13.1) bei einfacher Mittelbildung die Systeme, die einen von den Parametern unabhängigen Multiplikator gestatten, und diese sind, wie der Satz 6 behauptet, durch eine von den Parametern freie, also erlaubte Transformation in Liouvillesche Systeme überführbar.

Wir beweisen nunmehr diesen Satz 6. Führen wir zunächst statt der $x_1 \ldots x_n$ neue Koordinaten $\xi_1 \ldots \xi_n$ ein, die reguläre, umkehrbare, die Parameter $a_1 \ldots a_\varrho$ nicht enthaltende Funktionen der $x_1 \ldots x_n$ sind, so transformiert sich das System (1.1) in ein neues System

$$\frac{d\xi_i}{dt} = H_i = \sum_{\varkappa=1}^{n} \frac{\partial \xi_i}{\partial x_\varkappa} X_\varkappa \qquad (i = 1 \ldots n);$$

wir verlangen, daß dieses transformierte System vom Liouvilleschen Typus sei, also seine Divergenz verschwinde:

$$\sum_{i=1}^{n} \frac{\partial H_i}{\partial \xi_i} = \sum_{i=1}^{n}\sum_{\lambda=1}^{n} \frac{\partial H_i}{\partial x_\lambda}\frac{\partial x_\lambda}{\partial \xi_i} = 0$$

oder:

$$\sum_{i,\varkappa,\lambda=1}^{n} \frac{\partial^2 \xi_i}{\partial x_\varkappa\, \partial x_\lambda}\frac{\partial x_\lambda}{\partial \xi_i} X_\varkappa + \sum_{\varkappa,\lambda=1}^{n} \frac{\partial X_\varkappa}{\partial x_\lambda} \sum_{i=1}^{n} \frac{\partial x_\lambda}{\partial \xi_i}\frac{\partial \xi_i}{\partial x_\varkappa} = 0. \tag{13.3}$$

---

[17]) Vgl. Levi-Civita 1., S. 339–342.

Andrerseits gelten die Identitäten

$$\frac{\partial x_\lambda}{\partial x_\varkappa} = \sum_{i=1}^{n} \frac{\partial x_\lambda}{\partial \xi_i} \frac{\partial \xi_i}{\partial x_\varkappa} = \delta_{\lambda\varkappa},$$

aus denen man folgert:

$$\frac{\partial x_\lambda}{\partial \xi_i} \cdot \frac{\partial(\xi_1 \ldots \xi_n)}{\partial(x_1 \ldots x_n)} = (-1)^{i+\lambda} \frac{\partial(\xi_1 \ldots \xi_{i-1}\, \xi_{i+1} \ldots \xi_n)}{\partial(x_1 \ldots x_{\lambda-1}\, x_{\lambda+1} \ldots x_n)},$$

und daher erhält (13.3) die Form:

$$\sum_{\varkappa=1}^{n} X_\varkappa \sum_{i=1}^{n} \sum_{\lambda=1}^{n} (-1)^{i+\lambda} \frac{\partial^2 \xi_i}{\partial x_\varkappa \partial x_\lambda} \frac{\partial(\xi_1 \ldots \xi_{i-1}\, \xi_{i+1} \ldots \xi_n)}{\partial(x_1 \ldots x_{\lambda-1}\, x_{\lambda+1} \ldots x_n)} + \frac{\partial(\xi_1 \ldots \xi_n)}{\partial(x_1 \ldots x_n)} \sum_{\varkappa=1}^{n} \frac{\partial X_\varkappa}{\partial x_\varkappa} = 0$$

oder endlich:

$$(13.4) \qquad \sum_{\varkappa=1}^{n} \frac{\partial}{\partial x_\varkappa} \left\{ X_\varkappa \cdot \frac{\partial(\xi_1 \ldots \xi_n)}{\partial(x_1 \ldots x_n)} \right\} = 0,$$

woraus sich durch Vergleich mit (11.4) findet, daß die Funktionaldeterminante

$$(13.5) \qquad \frac{\partial(\xi_1 \ldots \xi_n)}{\partial(x_1 \ldots x_n)} = \mu$$

ein Multiplikator des Systems (1.1) sein muß. Die linke Seite ist nach Voraussetzung von den Parametern unabhängig, und das gleiche muß somit bezüglich eines der möglichen Multiplikatoren von (1.1) gelten.

Gibt es umgekehrt unter den Multiplikatoren von (1.1) einen — wir nennen ihn μ —, der von $a_1 \ldots a_\varrho$ unabhängig ist, so kann man stets eine Transformation der x in die ξ finden, deren Funktionaldeterminante (13.5) gleich μ, und die frei von den Parametern ist. Sie führt dann, obigen Gleichungen zufolge, das System (1.1) in ein solches von Liouvilleschem Charakter über. Damit ist der Satz 6 bewiesen.

§ 14.

Die Reduktion der imprimitiven Systeme.

Wir wenden uns nunmehr der Behandlung der allgemeinen m-fach imprimitiven Systeme, also in der Klassifikation des § 8 gesprochen, der Fälle A, B 2 und B 3 zu, für deren Invarianten die Differentialsysteme (10.13) und (11.10) charakteristisch sind. Wir kennen also die m unabhängigen Integrale von (1.1):

$$(14.1) \qquad f_i(x_1 \ldots x_n \,|\, a_1 \ldots a_\varrho) = c_i \qquad (i = 1 \ldots m \leqq n-1),$$

die im R_n die geschlossene Mannigfaltigkeit Φ bestimmen. Es ist der nächstliegende Gedanke, die Kenntnis dieser Integrale zur Erniedrigung des Ranges des ursprünglichen Differentialsystems nutzbar zu machen, um da-

durch auf ein einfach-imprimitives oder zweidimensionales System überzugehen.

Dazu führt folgendes Verfahren: wir nehmen die Integrale

$$f_1, f_2 \ldots f_{\varkappa-1}, f_{\varkappa+1} \ldots f_m \tag{14.2}$$

und lösen sie nach $m-1$ von den $x_1 \ldots x_n$ auf, was wegen ihrer Unabhängigkeit stets möglich ist. Diese Variablen, deren Wahl natürlich von $\varkappa$ abhängen kann, seien mit $x_1 \ldots x_{m-1}$ bezeichnet. Dann gibt also die Auflösung auf Φ das Bestehen von $m-1$ Beziehungen der Art:

$$x_i = \psi_{i\varkappa}(x_m \ldots x_n \,|\, a_1 \ldots a_\varrho \,|\, c_1 \ldots c_{\varkappa-1} c_{\varkappa+1} \ldots c_m) \qquad (i = 1 \ldots m-1), \tag{14.3}$$

die identisch den folgenden Gleichungen genügen:

$$f_\lambda(\psi_{1\varkappa} \ldots \psi_{m-1,\varkappa}, x_m \ldots x_n \,|\, a_1 \ldots a_\varrho) \equiv c_\lambda \\ (\lambda = 1 \ldots \varkappa - 1, \varkappa + 1 \ldots m). \tag{14.4}$$

Da die Größen (14.2) Integrale von (1.1) sind, folgen einmal die weiteren Identitäten:

$$\frac{d\psi_{i\varkappa}}{dt} = \sum_{j=m}^{n} \frac{\partial \psi_{i\varkappa}}{\partial x_j} X_j \equiv X_i \qquad (i = 1 \ldots m-1), \tag{14.5}$$

die man sich natürlich in den Variablen $x_m \ldots x_n$ geschrieben denken muß, andrerseits bleibt ein System von $n - m + 1$ Differentialgleichungen übrig:

$$\frac{dx_j}{dt} = X_j(\psi_{1\varkappa} \ldots \psi_{m-1,\varkappa}, x_m \ldots x_n \,|\, a_1 \ldots a_\varrho) \qquad (j = m \ldots n), \tag{14.6}$$

das nach unsern Voraussetzungen ein einziges bekanntes endlich-deutiges Integral zuläßt, nämlich

$$f_\varkappa(\psi_{1\varkappa} \ldots \psi_{m-1,\varkappa}, x_m \ldots x_n) = c_\varkappa, \tag{14.7}$$

also einfach-imprimitiv bzw., wenn $m = n - 1$, zweidimensional ist.

Dieses Integral ist nichts anderes als die durch Elimination der $x_1 \ldots x_{m-1}$ aus (14.1) hervorgehende Beziehung, d. h. stellt geometrisch gesprochen die Projektion Φ^* der Mannigfaltigkeit Φ aus dem R_n der $x_1 \ldots x_n$ auf den $n - m + 1$-dimensionalen Koordinatenraum R^* der $x_m \ldots x_n$ dar; woraus erhellt, daß diese geschlossene Mannigfaltigkeit Φ^* invariant ist gegenüber der Wahl des Index $\varkappa$, vorausgesetzt, daß die Variablen, nach denen aufgelöst wird, dabei die gleichen bleiben. Φ^* wird gemäß unsern topologischen Voraussetzungen von den Trajektorien des Systems (14.6) quasiergodisch bedeckt bzw. fällt mit diesen in der (x_{n-1}, x_n)-Ebene zusammen. Man wird nun statt des ursprünglichen Systems (1.1) diese reduzierten Systeme (14.6) untersuchen, und wir werden einmal die adiabatischen Invarianten der letzteren aufsuchen und dann nachsehen,

unter welchen Bedingungen diese auch Invarianten des ursprünglichen Systems sind.

Es wird dazu zweckmäßig sein, vorerst einige auf das System (14.6) bezügliche Beziehungen aufzustellen. Alle Größen können in unsern Rechnungen in zwei verschiedenen Formen auftreten: Einmal als Funktionen der $x_1 \dots x_n \mid a_1 \dots a_\varrho$ und frei von den $c_1 \dots c_m$, ein andermal als Funktionen der $x_m \dots x_n \mid a_1 \dots a_\varrho \mid c_1 \dots c_{k-1}\, c_{k+1} \dots c_m$; in dieser letzteren Form geschrieben, in der sie also die Projektion auf R^* darstellen, wollen wir sie mit einem Stern versehen. Ferner wird in unseren Ausführungen die Matrix

$$D \equiv \left(\frac{\partial f_i}{\partial x_1} \cdots \frac{\partial f_i}{\partial x_{m-1}}\right) \qquad (i = 1 \dots m) \tag{14.8}$$

eine Rolle spielen. Die durch Streichung der Zeilen $p, q, r \dots$ und der Spalten $\lambda, \mu \dots$ hervorgehenden Determinanten wollen wir durch

$$D_{p,\,q,\,r\,\dots}^{\lambda,\,\mu\,\dots}$$

bezeichnen.

Wir bestimmen zunächst die Divergenz von (14.6); es ist:

$$\varDelta_\varkappa = \sum_{j=m}^{n} \frac{\partial X_j^*}{\partial x_j} = \sum_{i=1}^{m-1} \sum_{j=m}^{n} \frac{\partial X_j}{\partial x_i} \frac{\partial \psi_{i\varkappa}}{\partial x_j} + \sum_{j=m}^{n} \frac{\partial X_j}{\partial x_j}. \tag{14.9}$$

Aus (14.4) folgen andrerseits die Gleichungen

$$\frac{\partial f_\lambda}{\partial x_1} \frac{\partial \psi_{1\varkappa}}{\partial x_j} + \dots + \frac{\partial f_\lambda}{\partial x_{m-1}} \frac{\partial \psi_{m-1,\varkappa}}{\partial x_j} = -\frac{\partial f_\lambda}{\partial x_j} \qquad (j = m \dots n;\ \lambda \neq \varkappa = 1 \dots m),$$

deren Auflösung ergibt:

$$\frac{\partial \psi_{i\varkappa}}{\partial x_j} = \frac{(-1)^{i+1}}{D_\varkappa} \left\{ \sum_{\lambda=1}^{\varkappa-1} (-1)^\lambda \frac{\partial f_\lambda}{\partial x_j} D_{\lambda\varkappa}^{i} - \sum_{\lambda=\varkappa+1}^{n} (-1)^\lambda \frac{\partial f_\lambda}{\partial x_j} D_{\varkappa\lambda}^{i} \right\}, \tag{14.10}$$

also

$$\varDelta_\varkappa = \frac{1}{D_\varkappa} \sum_{i=1}^{m-1} \sum_{j=m}^{n} (-1)^{i+1} \frac{\partial X_j}{\partial x_i} \left[\sum_{\lambda=1}^{\varkappa-1} (-1)^\lambda \frac{\partial f_\lambda}{\partial x_j} D_{\lambda\varkappa}^{i} - \sum_{\lambda=\varkappa+1}^{n} (-1)^\lambda \frac{\partial f_\lambda}{\partial x_j} D_{\varkappa\lambda}^{i} \right] + \sum_{i=m}^{n} \frac{\partial X_i}{\partial x_i}.$$

Nun sind aber die $f_1 \dots f_m$ Integrale von (1.1), d. h. es gelten die Identitäten

$$\sum_{j=1}^{n} \frac{\partial f_\lambda}{\partial x_j} X_j \equiv 0 \qquad (\lambda = 1 \dots m), \tag{14.11}$$

durch deren Differentiation nach x_i man findet:

$$\sum_{j=m}^{n} \frac{\partial f_\lambda}{\partial x_j} \frac{\partial X_j}{\partial x_i} = -\sum_{j=1}^{m-1} \frac{\partial f_\lambda}{\partial x_j} \frac{\partial X_j}{\partial x_i} - \sum_{j=1}^{n} \frac{\partial^2 f_\lambda}{\partial x_i\, \partial x_j} X_j,$$

so daß schließlich folgt:

$$\Delta_\varkappa = \frac{1}{D_\varkappa}\sum_{i=1}^{m-1}\sum_{j=1}^{m-1}\frac{\partial X_j}{\partial x_i}\left[\sum_{\lambda=1}^{\varkappa-1}(-1)^{i+\lambda}\frac{\partial f_\lambda}{\partial x_j}D^i_{\lambda\varkappa} - \sum_{\lambda=\varkappa+1}^{m}(-1)^{\lambda+i}\frac{\partial f_\lambda}{\partial x_j}D^i_{\varkappa\lambda}\right]$$

$$+\frac{1}{D_\varkappa}\sum_{j=1}^{n}X_j\sum_{i=1}^{m-1}\left[\sum_{\lambda=1}^{\varkappa-1}(-1)^{i+\lambda}\frac{\partial^2 f_\lambda}{\partial x_i\partial x_j}D^i_{\lambda\varkappa} - \sum_{\lambda=\varkappa+1}^{m}(-1)^{i+\lambda}\frac{\partial^2 f_\lambda}{\partial x_i\partial x_j}D^i_{\varkappa\lambda}\right] + \sum_{i=m}^{n}\frac{\partial X_i}{\partial x_i}$$

$$= \sum_{i=1}^{n}\frac{\partial X_i}{\partial x_i} + \sum_{j=1}^{n}X_j\frac{\partial \log D_\varkappa}{\partial x_j},$$

oder, wenn Δ_0 die Divergenz des ursprünglichen Systems bedeutet:

$$\Delta_\varkappa = \Delta_0 + \sum_{i=1}^{n}X_i\frac{\partial \log D_\varkappa}{\partial x_i}. \tag{14.12}$$

Für diese Ableitung ist die Bemerkung wichtig, daß $D_\varkappa$ nicht verschwinden kann, weil wir die Auflösbarkeit der Größen (14.2) nach den $x_1 \ldots x_{m-1}$ voraussetzen. Der Formel (14.12) entnimmt man beiläufig die Tatsache, daß die Divergenz des reduzierten Systems mit der des ursprünglichen identisch wird, wenn $D_\varkappa$ ein Integral von (1.1) ist, und speziell den

Satz 10. *Ist das ursprüngliche System von Liouvilleschem Typus, so gilt das gleiche vom reduzierten System dann und nur dann, wenn $D_\varkappa$ ein Integral des ersteren ist.*

Nunmehr sind wir in der Lage, den Multiplikator $\mu_\varkappa$ des reduzierten Systems (14.6) zu finden, dessen Differentialgleichung lautet:

$$\sum_{j=m}^{n}X_j^*\frac{\partial \mu_\varkappa^*}{\partial x_j} + \mu_\varkappa^*\Delta_\varkappa^* = 0,$$

oder

$$\sum_{j=m}^{n}\sum_{i=1}^{m-1}\frac{\partial \mu_\varkappa}{\partial x_i}\frac{\partial \psi_{i\varkappa}}{\partial x_j}X_j + \sum_{j=m}^{n}\frac{\partial \mu_\varkappa}{\partial x_j}X_j + \mu_\varkappa\Delta_\varkappa = 0,$$

was unter Benutzung von (14.10) und der aus (14.11) folgenden Identität

$$\sum_{j=m}^{n}\frac{\partial f_\lambda}{\partial x_j}X_j = -\sum_{j=1}^{m-1}\frac{\partial f_\lambda}{\partial x_j}X_j \qquad (\lambda = 1 \ldots m)$$

die Gestalt annimmt:

$$\sum_{i=1}^{m-1}\frac{1}{D_\varkappa}\frac{\partial \mu_\varkappa}{\partial x_i}\sum_{j=1}^{m-1}X_j\left\{\sum_{\lambda=1}^{\varkappa-1}(-1)^{i+\lambda}\frac{\partial f_\lambda}{\partial x_j}D^i_{\lambda\varkappa} - \sum_{\lambda=\varkappa+1}^{m}(-1)^{i+\lambda}\frac{\partial f_\lambda}{\partial x_j}D^i_{\varkappa\lambda}\right\}$$

$$+\sum_{j=m}^{n}\frac{\partial \mu_\varkappa}{\partial x_j}X_j + \mu_\varkappa\Delta_\varkappa = 0,$$

oder

$$\sum_{i=1}^{n} \frac{\partial \mu_\varkappa}{\partial x_i} X_i + \mu_\varkappa \Delta_\varkappa = 0; \tag{14.13}$$

in Verbindung mit (14.12) folgt hieraus:

$$\sum_{i=1}^{n} \frac{\partial \log(\mu_\varkappa D_\varkappa)}{\partial x_i} X_i + \Delta_0 = 0,$$

d. h. bezeichnet μ_0 einen Multiplikator des ursprünglichen Systems, so ist

$$\mu_\varkappa = \frac{\mu_0}{D_\varkappa} \tag{14.14}$$

ein solcher des reduzierten Systems (14.6).

Das System (14.6) ist — wenn wir ganz von seinem Ursprung absehen und es für sich allein betrachten — ein einfach-imprimitives $n-m+1$-dimensionales System, dessen rechte Seiten die Parameter $a_1 \ldots a_\varrho \,|\, c_1 \ldots c_{\varkappa-1}\, c_{\varkappa+1} \ldots c_m$ enthalten, und diese Parameter sind auf eine zweifache Weise hineingekommen: einmal waren die $a_1 \ldots a_\varrho$ schon in den X_j enthalten, sodann kommen sie, wie auch die c durch Ausführung der Substitution (14.3) noch einmal zum Vorschein. Wir wollen ihr Auftreten im Integral $f_\varkappa^*$ und Multiplikator $\mu_\varkappa^*$ eingehender untersuchen. Aus den Identitäten (14.4) folgt zunächst durch Differentiation nach c_λ

$$\frac{\partial f_j}{\partial x_1}\frac{\partial \psi_{1\varkappa}}{\partial c_\lambda} + \ldots + \frac{\partial f_j}{\partial x_{m-1}}\frac{\partial \psi_{m-1,\varkappa}}{\partial c_\lambda} \equiv \delta_{j\lambda} \qquad (j, \lambda = 1 \ldots \varkappa-1,\ \varkappa+1 \ldots m),$$

woraus sich durch Auflösung findet:

$$\frac{\partial \psi_{i\varkappa}}{\partial c_\lambda} = \begin{cases} (-1)^{i+\lambda}\,\dfrac{D^i_{\lambda\varkappa}}{D_\varkappa}, & \text{wenn } \lambda < \varkappa, \\[2ex] (-1)^{i+\lambda+1}\,\dfrac{D^i_{\varkappa\lambda}}{D_\varkappa}, & \text{wenn } \lambda > \varkappa. \end{cases} \tag{14.15}$$

Ferner folgt:

$$\frac{\partial f_\varkappa^*}{\partial c_\lambda} = \sum_{i=1}^{m-1} \frac{\partial f_\varkappa}{\partial x_i}\frac{\partial \psi_{i\varkappa}}{\partial c_\lambda} = (-1)^{\varkappa+\lambda+1}\frac{D_\lambda}{D_\varkappa} \qquad (\lambda \gtrless \varkappa) \tag{14.16}$$

und:

$$\frac{\partial \mu_\varkappa^*}{\partial c_\lambda} = \sum_{i=1}^{m-1} \frac{\partial \mu_\varkappa}{\partial x_i}\frac{\partial \psi_{i\varkappa}}{\partial c_\lambda} = \begin{cases} \dfrac{1}{D_\varkappa}\displaystyle\sum_{i=1}^{m-1}(-1)^{i+\lambda}\, D^i_{\lambda\varkappa}\frac{\partial \mu_\varkappa}{\partial x_i}, & \text{wenn } \lambda < \varkappa, \\[2ex] \dfrac{1}{D_\varkappa}\displaystyle\sum_{i=1}^{m-1}(-1)^{i+\lambda+1}\, D^i_{\varkappa\lambda}\frac{\partial \mu_\varkappa}{\partial x_i}, & \text{wenn } \lambda > \varkappa. \end{cases} \tag{14.17}$$

In ähnlicher Weise gibt die Differentiation von (14.4) nach a_ν die Gleichungen

$$\frac{\partial f_j}{\partial x_1}\frac{\partial \psi_{1\varkappa}}{\partial a_\nu} + \ldots + \frac{\partial f_j}{\partial x_{m-1}}\frac{\partial \psi_{m-1,\varkappa}}{\partial a_\nu} = -\frac{\partial f_j}{\partial a_\nu}$$
$$(j = 1 \ldots \varkappa-1,\ \varkappa+1 \ldots m;\ \nu = 1 \ldots \varrho),$$

aus denen folgt:

$$\frac{\partial \psi_{i\varkappa}}{\partial a_\nu} = \frac{1}{D_\varkappa}\left[-\sum_{\lambda=1}^{\varkappa-1}(-1)^{i+\lambda}\frac{\partial f_\lambda}{\partial a_\nu}D^i_{\lambda\varkappa} + \sum_{\lambda=\varkappa+1}^{m}(-1)^{i+\lambda}\frac{\partial f_\lambda}{\partial a_\nu}D^i_{\varkappa\lambda}\right]. \tag{14.18}$$

Mithin wird:

$$\frac{\partial f^*_\varkappa}{\partial a_\nu} = \sum_{i=1}^{m-1}\frac{\partial f_\varkappa}{\partial x_i}\frac{\partial \psi_{i\varkappa}}{\partial a_\nu} + \frac{\partial f_\varkappa}{\partial a_\nu} = \sum_{\lambda=1}^{\varkappa-1}(-1)^{\varkappa+\lambda}\frac{D_\lambda}{D_\varkappa}\frac{\partial f_\lambda}{\partial a_\nu} + \sum_{\lambda=\varkappa+1}^{m}(-1)^{\varkappa+\lambda}\frac{D_\lambda}{D_\varkappa}\frac{\partial f_\lambda}{\partial a_\nu} + \frac{\partial f_\varkappa}{\partial a_\nu}$$

$$= \sum_{\lambda=1}^{m}(-1)^{\varkappa+\lambda}\frac{\partial f_\lambda}{\partial a_\nu}\frac{D_\lambda}{D_\varkappa}. \tag{14.19}$$

Beiläufig folgen hieraus in Verbindung mit (14.16) die Formeln

$$\frac{\partial f^*_\varkappa}{\partial a_\nu} - \frac{\partial f_\varkappa}{\partial a_\nu} = -\sum_{\lambda=1}^{m}{}'\frac{\partial f_\lambda}{\partial a_\nu}\frac{\partial f^*_\varkappa}{\partial c_\lambda} \qquad (\nu = 1\ldots\varrho), \tag{14.20}$$

in denen der Akzent an der Summe das Auslassen des Gliedes $\lambda = \varkappa$ andeuten soll. Schließlich wird:

$$\frac{\partial \mu^*_\varkappa}{\partial a_\nu} = \sum_{i=1}^{m-1}\frac{\partial \mu_\varkappa}{\partial x_i}\frac{\partial \psi_{i\varkappa}}{\partial a_\nu} + \frac{\partial \mu_\varkappa}{\partial a_\nu} = \frac{1}{D_\varkappa}\left[\sum_{\lambda=1}^{\varkappa-1}\frac{\partial f_\lambda}{\partial a_\nu}\sum_{i=1}^{m-1}(-1)^{i+\lambda+1}\frac{\partial \mu_\varkappa}{\partial x_i}D^i_{\lambda\varkappa}\right.$$

$$\left. + \sum_{\lambda=\varkappa+1}^{m}\frac{\partial f_\lambda}{\partial a_\nu}\sum_{i=1}^{m-1}(-1)^{i+\lambda}\frac{\partial \mu_\varkappa}{\partial x_i}D^i_{\varkappa\lambda}\right] + \frac{\partial \mu_\varkappa}{\partial a_\nu},$$

$$\frac{\partial \mu^*_\varkappa}{\partial a_\nu} - \frac{\partial \mu_\varkappa}{\partial a_\nu} = -\sum_{\lambda=1}^{m}{}'\frac{\partial f_\lambda}{\partial a_\nu}\frac{\partial \mu^*_\varkappa}{\partial c_\lambda}. \tag{14.21}$$

§ 15.

Existenzbedingungen der adiabatischen Invarianten.

Das System (14.6) mit den Parametern $a_1\ldots a_\varrho$, $c_1\ldots c_{\varkappa-1}\,c_{\varkappa+1}\ldots c_m$ und dem Integral (14.7) kann eine adiabatische Invariante $J^*_\varkappa$ besitzen, die dann gemäß (9.11) durch das Gleichungssystem definiert ist:

$$\left\{\begin{aligned} &\frac{\partial J^*_\varkappa}{\partial c_\varkappa}\cdot\int_{\Phi^*}\frac{\partial f^*_\varkappa}{\partial c_\lambda}F^*\frac{\mu^*_\varkappa\,d\Phi^*}{\Gamma^*} + \frac{\partial J^*_\varkappa}{\partial c_\lambda}\cdot\int_{\Phi^*}F^*\frac{\mu^*_\varkappa\,d\Phi^*}{\Gamma^*} = 0 \\ &\hspace{14em}(\lambda = 1\ldots\varkappa-1,\ \varkappa+1\ldots m), \\ &\frac{\partial J^*_\varkappa}{\partial c_\varkappa}\cdot\int_{\Phi^*}\frac{\partial f^*_\varkappa}{\partial a_\nu}F^*\frac{\mu^*_\varkappa}{\Gamma^*}d\Phi^* + \frac{\partial J^*_\varkappa}{\partial a_\nu}\cdot\int_{\Phi^*}F^*\frac{\mu^*_\varkappa}{\Gamma^*}d\Phi^* = 0 \qquad (\nu = 1\ldots\varrho), \end{aligned}\right. \tag{15.1}$$

worin Γ^* das Analogon der durch (9.4) definierten Größe Γ_{1n} für $f^*_\varkappa$ bedeutet; $J^*_\varkappa$ ist eine bloße Funktion der $a_1\ldots a_\varrho$, $c_1\ldots c_m$. Die hier auftretenden Integrale über die Projektion Φ^* lassen sich leicht in solche

über Φ umformen. Man beachte, daß nach (10.10) das Raumelement des R_n die Größe

$$d\tau = dx_1 \ldots dx_n = \frac{1}{\Gamma_{mn}} d\Phi\, df_1\, df_2 \ldots df_m,$$

andrerseits nach (9.5) das Raumelement von R^* den Ausdruck

$$d\tau^* = dx_m \ldots dx_n = \frac{1}{\Gamma^*} d\Phi^*\, df_\varkappa$$

besitzt, woraus die Beziehungen — vgl. (14.14) —

$$\frac{d\Phi}{\Gamma_{m,n}} = \frac{1}{D_\varkappa}\frac{d\Phi^*}{\Gamma^*}, \qquad \mu_0 \frac{d\Phi}{\Gamma_{m,n}} = \mu_\varkappa \frac{d\Phi^*}{\Gamma^*} \tag{15.2}$$

folgen. Durch Übergang von der Projektion auf die Mannigfaltigkeit Φ können wir also den Gleichungen (15.1) die Gestalt geben:

$$\left\{\begin{aligned} &\frac{\partial J_\varkappa^*}{\partial c_\varkappa} \cdot \int\limits_\Phi \frac{\partial f_\varkappa^*}{\partial c_\lambda} F \frac{\mu_0}{\Gamma_{m,n}} d\Phi + \frac{\partial J_\varkappa^*}{\partial c_\lambda} \cdot \int\limits_\Phi F \frac{\mu_0}{\Gamma_{m,n}} d\Phi = 0 \\ &\qquad\qquad\qquad\qquad (\lambda = 1 \ldots \varkappa - 1, \varkappa + 1 \ldots m), \\ &\frac{\partial J_\varkappa^*}{\partial c_\varkappa} \cdot \int\limits_\Phi \frac{\partial f_\varkappa^*}{\partial a_\nu} F \frac{\mu_0}{\Gamma_{m,n}} d\Phi + \frac{\partial J_\varkappa^*}{\partial a_\nu} \cdot \int\limits_\Phi F \frac{\mu_0}{\Gamma_{m,n}} d\Phi = 0 \qquad (\nu = 1 \ldots \varrho). \end{aligned}\right. \tag{15.3}$$

Das reduzierte System war von uns eingeführt worden, um die Berechnung der adiabatischen Invarianten des ursprünglichen Systems (1.1) zu ermöglichen. A priori wäre man geneigt zu behaupten, daß jede adiabatische Invariante des reduzierten Systems eine solche des ursprünglichen ist; *wir werden jedoch sehen, daß dem nicht so ist,* daß dazu vielmehr gewisse Bedingungen erfüllt sein müssen. Unsere Frage lautet also: Wann ist eine Lösung von (15.3) auch eine solche von (10.13)? Da natürlich $\frac{\partial J_\varkappa^*}{\partial c_\varkappa} \neq 0$ ist, gibt die Elimination der Derivierten von $J_\varkappa^*$ aus (15.3) und (10.13) die Bedingungen:

$$\begin{aligned} &\sum_{\lambda=1}^{m} \int\limits_\Phi \frac{\partial f_\varkappa^*}{\partial c_\lambda} F \frac{\mu_0}{\Gamma_{m,n}} d\Phi \cdot \int\limits_\Phi \frac{\partial f_\lambda}{\partial a_\nu} F \frac{\mu_0}{\Gamma_{m,n}} d\Phi \\ &\qquad + \int\limits_\Phi \left\{ \frac{\partial f_\varkappa^*}{\partial a_\nu} - \frac{\partial f}{\partial a_\nu} \right\} F \frac{\mu_0}{\Gamma_{m,n}} d\Phi \cdot \int\limits_\Phi F \frac{\mu_0}{\Gamma_{m,n}} d\Phi = 0 \qquad (\nu = 1 \ldots \varrho), \end{aligned} \tag{15.4}$$

die mittels (10.12), (14.16) und (14.19) die einfache Form annehmen:

$$\sum_{\lambda=1}^{m} (-1)^\lambda \left\{ \overline{\frac{D_\lambda}{D_\varkappa} \frac{\partial f_\lambda}{\partial a_\nu}} - \overline{\frac{D_\lambda}{D_\varkappa}} \cdot \overline{\frac{\partial f_\lambda}{\partial a_\nu}} \right\} = 0 \qquad (\nu = 1 \ldots \varrho). \tag{15.5}$$

Satz 11. *Die notwendigen und hinreichenden Bedingungen dafür, daß eine adiabatische Invariante des reduzierten Systems auch eine solche des ursprünglichen sei, lauten:*

$$\sum_{\lambda=1}^{m}(-1)^{\lambda}\left\{\overline{\frac{D_\lambda}{D_\varkappa}\frac{\partial f_\lambda}{\partial a_\nu}}-\overline{\frac{D_\lambda}{D_\varkappa}}\cdot\overline{\frac{\partial f_\lambda}{\partial a_\nu}}\right\}=0 \qquad (\nu=1\ldots\varrho).$$

Beispielsweise ist (15.5) erfüllt, wenn gewisse unter den f_λ in der Sprechweise des § 12 stationäre Integrale und die ergänzenden $\frac{D_\lambda}{D_\varkappa}$ auf Φ konstant, also Funktionen von $f_1\ldots f_m$ sind. Der Fall, daß *sämtliche* $\frac{\partial f_\lambda}{\partial a_\nu}$ $(\lambda=1\ldots m,\ \nu=1\ldots\varrho)$ auf Φ konstant, also alle bekannten Integrale stationär gegenüber den Parametern sind, ist trivial, denn dann verschwinden die Mittelbildungen in (10.13) bzw. (1.8), J wird die *gewöhnliche* Invariante von (1.6), und für diese gibt es kein Analogon der Reduktionsbedingungen (15.5); eine andere Möglichkeit für das Erfülltsein von (15.5) ist die, daß sämtliche

$$\frac{D_\lambda}{D_\varkappa}=\Psi(f_1\ldots f_m\,|\,a_1\ldots a_\varrho) \qquad (\lambda=1\ldots\varkappa-1,\ \varkappa+1\ldots m) \tag{15.6}$$

sind; auch dies läßt sich deuten. Benutzt man nämlich zur Reduktion statt der Integrale (14.2) die folgenden: $f_1\ldots f_{\lambda-1},\ f_{\lambda+1}\ldots f_m$ und sind diese nach den gleichen Variablen $x_1\ldots x_{m-1}$ auflösbar, so sind die Multiplikatoren μ_λ $(\lambda=1\ldots m)$ sämtlicher m reduzierten Systeme nach (14.14) und (15.6) bis auf einen auf Φ konstanten Faktor miteinander identisch. Nach (14.13) wird dies im besonderen eintreten, *wenn die Divergenz Δ_λ für alle m reduzierten Systeme die gleiche ist.*

Ist also (15.5) erfüllt, so wissen wir, daß die adiabatische Invariante des reduzierten Systems auch eine solche des ursprünglichen ist, und wir haben demnach nur noch die Existenzbedingungen für erstere, d. h. die Integrabilitätsbedingungen von (15.1) hinzuzuschreiben. Dazu ziehen wir den Satz 7 heran und erhalten drei Gruppen von Bedingungen:

$$(15.7)\left\{\begin{aligned}
&\text{(a)}\quad \overline{\frac{\partial f_\varkappa^*}{\partial a_\nu}\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\mu}}-\overline{\frac{\partial f_\varkappa^*}{\partial a_\nu}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\mu}}=\overline{\frac{\partial f_\varkappa^*}{\partial a_\mu}\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\nu}}-\overline{\frac{\partial f_\varkappa^*}{\partial a_\mu}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\nu}}\\
&\hspace{22em}(\nu,\mu=1\ldots\varrho),\\
&\text{(b)}\quad \overline{\frac{\partial f_\varkappa^*}{\partial c_\lambda}\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\nu}}-\overline{\frac{\partial f_\varkappa^*}{\partial c_\lambda}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial a_\nu}}=\overline{\frac{\partial f_\varkappa^*}{\partial a_\nu}\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_\lambda}}-\overline{\frac{\partial f_\varkappa^*}{\partial a_\nu}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_\lambda}}\\
&\hspace{18em}(\lambda=1\ldots\varkappa-1,\ \varkappa+1\ldots m),\\
&\text{(c)}\quad \overline{\frac{\partial f_\varkappa^*}{\partial c_\lambda}\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_j}}-\overline{\frac{\partial f_\varkappa^*}{\partial c_\lambda}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_j}}=\overline{\frac{\partial f_\varkappa^*}{\partial c_j}\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_\lambda}}-\overline{\frac{\partial f_\varkappa^*}{\partial c_j}}\cdot\overline{\frac{\partial\log\mu_\varkappa^*F^*}{\partial c_\lambda}}\\
&\hspace{17em}(j,\lambda=1\ldots\varkappa-1,\ \varkappa+1\ldots m),
\end{aligned}\right.$$

worin wir nach den vorangehenden Bemerkungen die Mittelbildung im Sinne von (10.12) über Φ erstreckt denken können. Mittels (14.16) bis (14.20) und (15.5) lassen sie sich auf eine, allerdings wenig übersichtliche Form bringen, in der das Auftreten der gesternten Größen vermieden wird. Sind die $\frac{1}{2}\varrho(\varrho-1)+(m-1)\varrho+\frac{1}{2}(m-1)(m-2)$ Bedingungen (15.7), sowie die ϱ Gleichungen (15.5) erfüllt, so wissen wir nach vorangehendem, daß das reduzierte System (14.6) eine adiabatische Invariante besitzt und daß diese zugleich eine solche des ursprünglichen Systems (1.1) ist. Die Gesamtheit dieser Bedingungen ist stärker als die der Integrabilitätsbedingungen von (10.13), die in extenso eine sehr komplizierte Form haben[18]), bietet aber den Vorteil übersichtlicher Handhabung.

Zwei Spezialfälle, in denen die Beziehungen (15.7) befriedigt sind, sind erwähnenswert:

$$\text{(I)}\quad \frac{\partial \log \mu_\varkappa^* F^*}{\partial c_j} = A_j(f_1^* \ldots f_m^* \,|\, a_1 \ldots a_\varrho); \quad \frac{\partial \log \mu_\varkappa^* F^*}{\partial a_\nu} = B_\nu(f_1^* \ldots f_m^* \,|\, a_1 \ldots a_\varrho)$$
$$(j = 1 \ldots \varkappa-1,\ \varkappa+1 \ldots m;\ \nu = 1 \ldots \varrho),$$

d. h. die linken Seiten sind auf Φ konstant;

$$\text{(II)}\quad \frac{\partial f_\varkappa^*}{\partial c_j} = \bar{A}_j(f_1^* \ldots f_m^* \,|\, a_1 \ldots a_\varrho); \quad \frac{\partial f_\varkappa^*}{\partial a_\nu} = \bar{B}_\nu(f_1^* \ldots f_m^* \,|\, a_1 \ldots a_\varrho).$$

Der Fall der Proportionalität der linken Seiten in (I), (II) läßt sich wie in § 12 auf die beiden eben genannten Möglichkeiten reduzieren.

Fall I. Ist $F \equiv 1$, so ist, in früherer Ausdrucksweise gesprochen, $\mu_\varkappa$ ein stationärer Multiplikator des Systems (14.6). Sind sämtliche $A_j = 0$, so ist also

$$(15.8)\qquad \frac{\partial \mu_\varkappa^* F^*}{\partial c_j} = 0 \qquad (j = 1 \ldots \varkappa-1,\ \varkappa+1 \ldots m),$$

und das besagt nach (14.17) — denn eine ähnliche Gleichung wie jene gilt auch für $\mu_\varkappa^* F^*$ — daß

$$(15.9)\qquad \frac{\partial \mu_\varkappa F}{\partial x_i} = 0 \qquad (i = 1 \ldots m-1)$$

gelten muß. (14.21) gibt dann weiter

$$\frac{\partial \mu_\varkappa^* F^*}{\partial a_\nu} = \frac{\partial \mu_\varkappa F}{\partial a_\nu} \qquad (\nu = 1 \ldots \varrho),$$

so daß also

$$(15.10)\qquad \frac{\partial \log \mu_\varkappa F}{\partial a_\nu} = B_\nu(f_1 \ldots f_m) \qquad (\nu = 1 \ldots \varrho)$$

sein muß. Sind auch alle $B_\nu = 0$, so wird demnach

$$(15.11)\qquad \frac{\partial \mu_\varkappa F}{\partial a_\nu} = 0 \qquad (\nu = 1 \ldots \varrho),$$

[18]) In dieser Beziehung ist zu erwähnen, daß Fermi 1 die Aufstellung dieser Bedingungen für Hamiltonsche Systeme versucht; er beweist nur ihre Existenz an einem Beispiel.

was in Verbindung mit (15.9) besagt, daß *die Größe*

$$\mu_\varkappa F = \frac{\mu_0}{D_\varkappa} F$$

von $x_1 \ldots x_{m-1}$, $a_1 \ldots a_\varrho$ *nicht abhängen darf.* In früherer Ausdrucksweise ist dann das System (14.6) von generalisiert-Liouvilleschem Typus. Hinzu tritt dann noch die Reduktionsbedingung (15.5); hat diese die besondere Form (15.6), so sind alle möglichen reduzierten Systeme (wenn $D_\lambda \neq 0$ ist) vom generalisiert-Liouvilleschem Typus.

Fall II. $f_\varkappa^*$ ist ein stationäres Integral des Differentialsystems (14.6), und aus den Gleichungen (15.1) verschwindet die Funktionaloperation. Dieser Fall tritt im besonderen ein, wenn sämtliche $\bar{A}_j$, $\bar{B}_\nu$ verschwinden, d. h. nach (14.16) und (14.20), wenn

$$D_\lambda = 0 \qquad (\lambda \neq \varkappa = 1 \ldots m) \tag{15.12}$$

und

$$\frac{\partial f_\varkappa^*}{\partial a_\nu} = \frac{\partial f_\varkappa}{\partial a_\nu} = 0 \qquad (\nu = 1 \ldots \varrho), \tag{15.13}$$

d. h. *wenn* $f_\varkappa$ *weder von* $x_1 \ldots x_{m-1}$, *noch von* $a_1 \ldots a_\varrho$ *abhängt.* Wegen (15.12) sind dann die Reduktionsbedingungen (15.5) identisch befriedigt. $f_\varkappa$ ist dann also auch ein stationäres Integral des ursprünglichen Systems.

Tritt die oben erwähnte Möglichkeit des gleichzeitigen Bestehens von (15.9) und (15.11) ein, so gibt uns der Satz 9 direkt die adiabatische Invariante des reduzierten, und damit des ursprünglichen Systems

$$\begin{aligned} J = J_\varkappa^* &= \int\limits_{f_\varkappa^*} \mu_\varkappa^* F^* \, dx_m \ldots dx_n = \int\limits_{f_\varkappa^*} \frac{\mu_0^*}{D_\varkappa^*} F^* \, dx_m \ldots dx_n \\ &= \int\limits_{f_\varkappa^*} \frac{\mu_0}{D_\varkappa} F \, dx_m \ldots dx_n . \end{aligned} \tag{15.14}$$

Es ist günstigstenfalls möglich, durch Ausführung der m vorhandenen Reduktionen auf diese Weise sämtliche m verschiedenen Invarianten zu berechnen, die die $c_1 \ldots c_m$ mit den $a_1 \ldots a_\varrho$ in Verbindung setzen; dies wird nämlich dann der Fall sein, *wenn sämtliche reduzierten Systeme vom generalisiert-Liouvilleschen Typus mit wesentlich verschiedenen Multiplikatoren sind.* In den Fällen A und B 3 der Klassifikation des § 8 ist entsprechend der Satz 5 anzuwenden. Wir fassen zusammen:

Satz 12. *Außer der notwendigen Reduktionsbedingung* (15.15) *sind hinreichende Bedingungen für die Existenz einer adiabatischen Invariante eines m-fach imprimitiven Systems* a) *daß* $\frac{\mu_0 F}{D_\varkappa}$ *oder* b) *daß* $f_\varkappa$ *von*

$x_1 \ldots x_{m-1}, a_1 \ldots a_\varrho$ *nicht abhängt. Im ersten Falle ist*

$$J = \int_{\Phi^*} \frac{\mu_0 F}{D_\varkappa} dx_m \ldots dx_n$$

eine adiabatische Invariante.

Dieser Satz gestattet viele Anwendungen, wir wollen nur einen Spezialfall erwähnen:

Satz 13. *Sind das ursprüngliche, wie alle m reduzierten Systeme vom Liouvilleschen Typus und* $D_\varkappa \neq 0$ $(\varkappa = 1 \ldots m)$, *so ist der Rauminhalt der Projektion*

$$J = \int_{\Phi^*} dx_m \ldots dx_n$$

bei einfacher Mittelbildung adiabatisch invariant.

Denn dann sind alle $\mu_\varkappa$ auf Φ konstant, das gleiche gilt von μ_0, mithin von $D_\varkappa$, und (15.5) ist identisch erfüllt.

§ 16.
Ein Beispiel.

Als Beispiel, das die vorangehende Theorie illustriert, wollen wir die durch das System

$$(16.1)\quad \frac{dx_1}{dt} = a_2 x_2 - a_3 x_3; \quad \frac{dx_2}{dt} = a_3 x_3 - a_1 x_1; \quad \frac{dx_3}{dt} = a_1 x_1 - a_2 x_2$$

charakterisierte ebene elliptische Schwingung behandeln. Das System gestattet die beiden Integrale:

$$f_1 \equiv x_1 + x_2 + x_3 = c_1, \quad f_2 \equiv a_1 x_1^2 + a_2 x_2^2 + a_3 x_3^2 = c_2,$$

wir befinden uns somit im Falle B 3. Wir nehmen einfache Mittelbildung $F \equiv 1$ und $T = \infty$. Zur Berechnung der adiabatischen Invarianten stehen uns hier zwei Wege zur Verfügung: 1. die direkte Methode des § 1 unter Benutzung der x_i als Funktionen von $t - t_0$, und 2. das Verfahren des § 15 mittels Reduktion auf eine zweidimensionale Koordinatenebene.

Ad 1. Setzt man zur Abkürzung

$$h^2 = a_1 a_2 + a_2 a_3 + a_3 a_1, \quad k^2 = c_2 h^2 - a_1 a_2 a_3 c_1^2,$$

so gibt die Integration von (16.1) die Beziehungen

$$x_1 = \frac{1}{h^2} \frac{k}{\sqrt{a_1 + a_2}} \left\{ a_2 \cos h (t - t_0) + h \sin h (t - t_0) \right\} + \frac{a_2 a_3 c_1}{h^2},$$

$$x_2 = \frac{1}{h^2} \frac{k}{\sqrt{a_1 + a_2}} \left\{ a_1 \cos h (t - t_0) - h \sin h (t - t_0) \right\} + \frac{a_1 a_3 c_1}{h^2},$$

$$x_3 = -\frac{1}{h^2} k \sqrt{a_1 + a_2} \cos h (t - t_0) + \frac{a_1 a_2 c_1}{h^2}.$$

In der Bezeichnung (1.5) wird somit:

$$\varphi_{11} = \varphi_{12} = \varphi_{13} = 0,$$

$$\bar{\varphi}_{21} = \frac{1}{2h^4}\{2a_2^2 a_3^2 c_1^2 + k^2(a_2 + a_3)\}, \qquad \bar{\varphi}_{22} = \frac{1}{2h^4}\{2a_1^2 a_3^2 c_1^2 + k^2(a_1 + a_3)\},$$

$$\bar{\varphi}_{23} = \frac{1}{2h^4}\{2a_1^2 a_2^2 c_1^2 + k^2(a_1 + a_2)\}.$$

Damit ist das Gleichungssystem (1.11) aufzustellen, und man verifiziert leicht, daß seine beiden Integrale lauten

$$J_1 = c_1; \quad J_2 = \frac{k^2}{h^3}. \tag{16.2}$$

Ad 2. Zunächst ist nach Satz 1 c_1 eine adiabatische Invariante, also

$$J_1 = c_1.$$

Zur Reduktion verwenden wir f_1, d. h. setzen

$$x_1 \equiv \psi_{12} = c_1 - x_2 - x_3,$$

so daß sich (16.1) auf das folgende System reduziert:

$$\begin{aligned} \frac{dx_2}{dt} &= a_1 x_2 + (a_1 + a_3)x_3 - a_1 c_1; \\ \frac{dx_3}{dt} &= -(a_1 + a_2)x_2 - a_1 x_3 + a_1 c_1 \end{aligned} \tag{16.3}$$

mit dem Integral

$$\begin{aligned} \Phi^* = f_2^* \equiv (a_1 + a_2)x_2^2 + (a_1 + a_3)x_3^2 + 2a_1 x_2 x_3 \\ - 2a_1 c_1 (x_2 + x_3) + a_1 c_1^2 - c_2 = 0. \end{aligned} \tag{16.4}$$

Sind a_1, a_2, a_3 positiv, so ist die Trajektorie im R_3 der Schnitt des Ellipsoids f_2 mit der festen Ebene f_1, also eine Ellipse; Φ^* ist deren Projektion auf die x_2, x_3-Ebene, also auch eine Ellipse. Das reduzierte System (16.3) ist, wie das ursprüngliche (16.1), von der Divergenz Null, also

$$\mu_0 = \mu_2 = 1.$$

Die Reduktionsbedingung (15.5) ist wegen $\frac{\partial f_1}{\partial a_\nu} = 0$ erfüllt, desgleichen nach Satz 12 die Integrabilitätsbedingungen. Die Invariante ist der Flächeninhalt der Ellipse Φ^*, der sich zu

$$J_2 = \pi \frac{k^2}{h^3}$$

berechnet. Die Reduktion mittels f_2 hingegen ist nicht anwendbar.

Literaturnachweis.

Born, Max. 1. Vorlesungen über Atomdynamik. Bd. I. Berlin 1925.

Burgers, J. M. 1. Philos. Mag. **33** (1917), S. 514—520,
2. Dissertation Univers. Leiden. Haarlem 1918.

Carathéodory, Constantin. 1. Über den Wiederkehrsatz von Poincaré. Sitzungsber. d. Preuß. Akad. d. Wiss. 1919, S. 580—584.

Cherry, T. M. 1. Transactions of the Cambridge Philosophical Society **23** (1924), S. 43f.

Ehrenfest, Paul. 1. Adiabatic invariants and the theory of quanta. Phil. Mag. **33** (1917), S. 500—513.

Fermi, Enrico. 1. Alcuni teoremi di meccanica analitica. Nuovo Cimento (7) **25** (1923).

Geppert, Harald. 1. Sugli invarianti adiabatici di un generico sistema differenziale. Nota I. Rend. Acc. Lincei **8**, 2º sem. 1928, S. 30—34.
2. Idem, Nota II, ibidem S. 191—198.
3. Idem, Nota III, ibidem S. 294—299.
4. Die adiabatischen Invarianten beliebiger Differentialsysteme. Atti Congresso Intern. dei Matematici, Bologna 1928.

Gibbs, William. 1. Statistical mechanics. Yale Univers. Press 1902.

Hertz, Paul. 1. Weber-Gans, Repertorium d. Physik, Bd. I, 2. Leipzig 1916.

Levi-Civita, Tullio. 1. Drei Vorlesungen über adiabatische Invarianten. Abh. d. Math. Seminars Hamburg **6** (1928), S. 323—366.
2. Sugli invarianti adiabatici. Atti Congresso Intern. dei Fisici **2** (1927), S. 1—39.
3. Alcune applicazioni astronomiche degli invarianti adiabatici. Bollettino dell' Unione Mat. Italiana 7, Nr. 4, 1928.

Poincaré, Henri. 1. Les méthodes nouvelles de la mécanique céleste. Bd. 3. Paris 1899.

Smekal, Adolf. 1. Statistische und molekulare Theorie der Wärme, in Handbuch der Physik, Bd. 9. Berlin 1927.

(Eingegangen am 5. 12. 1928.)

Zur Geometrie der kontinuierlichen Transformationsgruppen.

Von

J. A. Schouten in Delft.

Einleitung.

Den beiden großen Abschnitten der Theorie der kontinuierlichen Gruppen, „Klassifizierungstheorie“ und „Darstellungstheorie“, ist kürzlich ein neuer Abschnitt zur Seite getreten, der methodisch den beiden anderen voranzugehen hat und geeignet wäre, die klassische Theorie, sofern sie nicht den beiden anderen Abschnitten angehört, in sich aufzunehmen. Dieser Abschnitt, den ich mit dem Namen „Geometrie der Gruppenmannigfaltigkeit“ bezeichnen möchte, zerfällt in zwei grundverschiedene Teile, Geometrie des Gruppenkeims, im kleinen, und Geometrie des Zusammenhangs, im großen. Vom ersten Teil soll im folgenden eine Übersicht des Wichtigsten gegeben werden, eine Übersicht, erwachsen aus Vorlesungen, die ich 1926—1927 an der Universität zu Leiden zu halten die Ehre hatte als Vorbereitung der Vorlesungen in 1927—1928 über Klassifizierungstheorie. In dieser Übersicht werden natürlich die Resultate der klassischen Theorie, insbesondere die drei Fundamentaltheoreme, vorausgesetzt. Das erste Kapitel bringt eine vorläufige synthetische Orientierung. Das zweite Kapitel zeigt, wie sich die drei Übertragungen und die wichtigsten Größen der Geometrie des Gruppenkeims auf kürzestem Wege aus den drei Fundamentaltheoremen ableiten. Im dritten Kapitel wird die Bedeutung der adjungierten Gruppe und ihrer Komitanten für die Geometrie des Gruppenkeims skizziert. Einer geschlossenen Darstellung zuliebe konnte es nicht immer umgangen werden, hin und wieder ein Resultat der klassischen Theorie mit zutage zu fördern. Dagegen sind Übergriffe in das Gebiet der Klassifizierungstheorie (bis auf die bloße Angabe der Kriterien für halbeinfache und integrabele Gruppen) streng vermieden und führt die Arbeit eben bis zu dem Punkte, wo diese Theorie einzusetzen hat.

I. Vorläufige Orientierung; Synthetisches.

§ 1.
Allgemeines.

Es seien

$$(1)\qquad {}'\overset{z}{x} = \overset{z}{f}(\overset{y}{x}, \overset{\nu}{\xi}) \qquad (y, z = 1, \ldots, n;\ \alpha, \ldots, \omega = a_1, \ldots, a_r)$$

die Gleichungen einer kontinuierlichen Transformationsgruppe in n Variablen mit r wesentlichen Parametern. Die ξ lassen sich als Koordinaten einer X_r auffassen. Wir betrachten im folgenden nur den *Gruppenkeim*, d. i. eine Umgebung der identischen Transformationen $\underset{0}{\xi}$, so klein, daß in ihr die $\overset{z}{f}$ regulär sind und eine eindeutige Korrespondenz zwischen den Punkten der X_r und den Transformationen der Gruppe besteht. Wo also von X_r gesprochen wird, ist nur der Teil von X_r gemeint, der mit dem Gruppenkeim korrespondiert.

Transformationen sollen mit großen Buchstaben angegeben werden, z. B. S, T, U, die Identität insbesondere mit I. Sollen die Parameter mit angedeutet werden, so schreiben wir T_ξ für die Transformation mit den Parametern ξ. Unter T^{-1} verstehen wir die Umkehrung von T, unter $T_\eta T_\xi$ die Transformation, die entsteht, indem *zuerst* T_ξ und *dann* T_η wirkt. Dabei gehen die x vermöge T_ξ über in x', und T_η ist also auf die x' anzuwenden:

$$(2)\qquad T_\eta T_\xi:\quad \overset{z}{x}'' = \overset{z}{f}(x', \eta) = \overset{z}{f}\{\overset{z}{f}(x, \xi), \eta\}.$$

Aus (1) folgt, daß man dieselbe Transformation erhält, wenn *zuerst* T_η auf die x wirkt und *dann* T_ξ auf die x in $\overset{z}{f}(x, \eta)$ angewandt wird.

§ 2.
Die Äquipollenzen[1]).

Zwei geordnete Punktepaare (S, T) und (A, B) in X_r sollen $(+)$-*äquipollent* bzw. $(-)$-*äquipollent* heißen, wenn

$$(3)\qquad T S^{-1} = B A^{-1} \quad \text{bzw.} \quad S^{-1} T = A^{-1} B.$$

Die Äquipollenzen haben u. a. folgende unmittelbar aus der Definition folgenden Eigenschaften[2]):

---

[1]) E. Cartan und J. A. Schouten, On the geometry of the group-manifold of simple and semi-simple groups, Proc. Kon. Akad. v. W. **29** (1926), S. 803–815, hier weiter angedeutet mit C & S 1.

[2]) Eine vollständige Aufzählung findet man bei Cartan, La géométrie des groupes de transformations, Journal de Math. p. et. a. **6** (1927), S. 1–119, hier weiter angedeutet mit C. G.

1. S, T und A bestimmen B eindeutig; S, T und B bestimmen A eindeutig.

2. Transitivität: aus (S, T) $(\pm)$-äquipollent (A, B) $(\pm)$-äquipollent (U, V) folgt (S, T) $(\pm)$-äquipollent (U, V).

3. Aus (S, T) $(\pm)$-äquipollent (A, B); (T, U) $(\pm)$-äquipollent (B, C) folgt (S, U) $(\pm)$-äquipollent (A, C).

4. Wenn von vier Punkten A, B, C, D die Paare (A, B) und (C, D) $(\pm)$-äquipollent sind, so sind die Paare (A, C) und (B, D) $(\mp)$-äquipollent.

Aus 1. folgt: Zu jedem Punkte S gibt es eine endliche Umgebung, die durch die $(\pm)$-Äquipollenz eineindeutig auf eine Umgebung jedes anderen Punktes T abgebildet wird. Läßt man T mit der identischen Transformation zusammenfallen, so folgt überdies, daß die $(\pm)$-Äquipollenz jedem geordneten Punktepaar in eineindeutiger Weise eine Transformation zuordnet. Jedes gerichtete Linienelement $(T_{\xi}, T_{\xi+d\xi})$ wird vermittels der $(\pm)$-Äquipollenz abgebildet auf eine *infinitesimale* Transformation:

$$
(4) \qquad \begin{cases} T_{\xi+d\xi}\, T_{\xi}^{-1} = T_{\underset{0}{\xi}+\delta_1\xi}\, T_{\underset{0}{\xi}}^{-1} = \bar{T}_{\underset{0}{\xi}+\delta_1\xi}\,, \\ T_{\xi}^{-1}\, T_{\xi+d\xi} = T_{\underset{0}{\xi}}^{-1}\, T_{\underset{0}{\xi}+\delta_2\xi} = T_{\xi+\delta_2\xi}\,. \end{cases}
$$

Im allgemeinen ist $\delta_1\xi \neq \delta_2\xi$.

Aus 2. folgt: Die solcherweise zustande kommenden zwei Abbildungen der Umgebungen der Punkte von X_r aufeinander sind eine jede unabhängig von der Wahl von S.

Aus 3. folgt: Die unendlich kleinen Umgebungen erster Ordnung werden durch jede der beiden Äquipollenzen *affin* aufeinander abgebildet. Die X_r trägt also *zwei* im allgemeinen verschiedene euklidische (d. h. integrabele) *lineare Übertragungen*, die $(+)$- und die $(-)$-Übertragung. Der Begriff der $(\pm)$-Äquipollenz überträgt sich so auf *Vektoren*. Ein Feld, das bei der $(\pm)$-Übertragung invariant ist, soll $(\pm)$-*konstant* heißen. Vermöge der $(\pm)$-Äquipollenz zugeordnete Richtungen heißen $(\pm)$-parallel. Die nähere analytische Bestimmung erfolgt später aus den Fundamentaltheoremen.

Eine Kurve heißt *geodätisch in X_r*, wenn sie zu je drei Punkten S, T und U stets auch $TS^{-1}U$ enthält. Daraus folgt, daß eine geodätische Linie entsteht, indem ein Linienelement stets entweder $(+)$- oder auch $(-)$-parallel in der eigenen Richtung verschoben wird. Eine geodätische Linie ist also in sich sowohl $(+)$- als $(-)$-parallel, aber zwei geodätische Linien, die $(\pm)$-parallel sind, sind im allgemeinen nicht $(\mp)$-parallel. Aus 1. folgt, daß durch zwei Punkte nur eine geodätische Linie geht; dies liegt natürlich daran, daß wir uns ausdrücklich auf den Gruppenkeim beschränken.

§ 3.

Die Parametergruppen und die adjungierte Gruppe.

Ist T eine allgemeine Transformation der Gruppe und durchläuft U sämtliche Transformationen der Gruppe, so stellt bekanntlich

(5) $$'T = U\,T \quad \text{bzw.} \quad 'T = T\,U$$

eine r-gliedrige Gruppe von Transformationen der Punkte der X_r dar, die *erste* bzw. *zweite Parametergruppe*. Aus der Definition der Äquipollenz folgt[3]):

1. Ein Punktepaar geht vermöge einer Transformation der ersten bzw. zweiten Parametergruppe über in ein $(-)$- bzw. $(+)$-äquipollentes Punktepaar.

2. $(+)$- und $(-)$-Äquipollenz bleiben bei beiden Gruppen erhalten.

3. Die Punktepaare (T, T') sind alle $(+)$- bzw. $(-)$-äquipollent.

Die beiden Parametergruppen sind unter sich und mit der gegebenen Gruppe *holoedrisch isomorph*; zugeordnet sind U, $'T = U\,T$ und $'T = T\,U^{-1}$.

Zur geometrischen Deutung der Parametergruppen bemerken wir, daß die Gleichung (1) zwei Deutungen zuläßt:

1. Als *Punkttransformation* der X_n, wobei der *Punkt* mit den Koordinaten x *übergeht* in den Punkt mit den Koordinaten $'x$ in bezug auf *dasselbe* Koordinatensystem.

2. Als *Koordinatentransformation*, wobei der *Punkt fest bleibt*, aber in bezug auf das *neue* Koordinatensystem die Koordinaten $'x$ bekommt.

Es sei nun das Koordinatensystem $\overset{s}{x}$, wie üblich, durch n Systeme von ∞^1 $(n-1)$-dimensionalen Mannigfaltigkeiten, von denen einer jeden eine Zahl zugeordnet ist, veranschaulicht. Ein solches veranschaulichtes Koordinatensystem wollen wir das „Gitter der x“ nennen und mit Σ_0[4]) bezeichnen. Dieses Gitter hat also nichts mit dem aus lauter Punkten bestehenden Gitter der Zahlentheorie zu tun.

Das „Gitter der $'x$“, d. i. das Gitter, in bezug auf welches dieselben festgebliebenen Punkte die Koordinaten $'x$ haben, wollen wir Σ_ξ nennen. Es entsteht, indem auf Σ_0 die Punkttransformation T_ξ^{-1} angewandt wird. Wir schreiben dies

(6) $$\Sigma_\xi = \Sigma_0\,T_\xi^{-1}.$$

---

[3]) C & S I, S. 812.

[4]) E. Cartan verwendet C. G., S. 19 für Σ_0 eine beliebige Figur, die so gewählt ist, daß eine eineindeutige Korrespondenz zwischen den transformierten Figuren und den Transformationen besteht. Bei ihm entsteht Σ_ξ durch T_ξ, nicht durch T_ξ^{-1}. Die hier verwendete Methode bietet den Vorteil, daß keine neuen Figuren eingeführt zu werden brauchen, da man mit den doch einmal vorhandenen Koordinatensystemen auskommt, und *daß $'x = T_\xi\,x$ gerade die Koordinaten in bezug auf Σ_ξ werden.*

Änderung des Gitters vermöge T_ξ^{-1} ist also gleichbedeutend mit Änderung der Koordinaten vermöge T_ξ.

Die Gruppenmannigfaltigkeit ist jetzt eineindeutig abgebildet auf die Gesamtheit aller Gitter Σ und es entspricht dabei Σ_ξ der Transformation T_ξ.

Sei jetzt $\bar{x} = Tx$ eine beliebige Transformation der x, so kann man sich die Frage stellen, wie diese Punkttransformation geschrieben wird in bezug auf das Gitter Σ_ξ. In bezug auf dieses Gitter sind die Koordinaten vor und nach der Transformation $'x = T_\xi x$ bzw. $'\bar{x} = T_\xi \bar{x}$, und es ist also

$$(7) \qquad '\bar{x} = T_\xi \bar{x} = T_\xi T x = T_\xi T T_\xi^{-1}\,'x.$$

Es ist also dasselbe, ob T ausgeführt wird auf die x oder $T_\xi T T_\xi^{-1}$ auf die $'x = T_\xi x$. Da auch $T_\xi T T_\xi^{-1}$ eine Transformation der Gruppe ist, werden die Transformationen der Gruppe beim Übergang von Σ_0 zu Σ_ξ vertauscht. Läßt man Σ_ξ nacheinander alle möglichen Stellungen einnehmen, so bilden diese Vertauschungen bekanntlich eine Gruppe, die *adjungierte* Gruppe. Da in der X_r jede Transformation mit einem Gitter korrespondiert, ist die adjungierte Gruppe eine Gruppe von Vertauschungen dieser Gitter. Die mit T_ξ korrespondierende Transformation führt dabei das zu T gehörige Bezugssystem $\Sigma = \Sigma_0 T^{-1}$ über in $\Sigma_0 T_\xi T^{-1} T_\xi^{-1}$.

Betrachten wir zwei Gitter Σ_ξ und Σ_η mit den zugehörigen Koordinaten $'x$ und $''x$. Dann geht Σ_ξ über in Σ_η durch die Transformation $T_\eta^{-1} T_\xi$ angewandt auf die x, und dies ist, wie wir sahen, gleichwertig mit $T_\xi T_\eta^{-1}$ wirkend auf die $'x$. $T_\eta^{-1} T_\xi$ gibt also die „absolute" gegenseitige Lage von Σ_ξ und Σ_η an, d. h. die Lage in bezug auf Σ_0, $T_\xi T_\eta^{-1}$ dagegen die „relative" gegenseitige Lage, d. h. die Lage in bezug auf Σ_ξ. Die gegenseitige Lage in bezug auf Σ_η ist gegeben durch $T_\eta T_\xi^{-1} = (T_\xi T_\eta^{-1})^{-1}$, und es entsteht dabei also nichts wesentlich Neues. Aus der Definition der $(+)$-Äquipollenz folgt also, daß zwei Gitterpaare $(\Sigma_\xi, \Sigma_\eta)$ und $(\Sigma_{'\xi}, \Sigma_{'\eta})$ $(+)$-äquipollent sind, wenn sie gleiche *relative* gegenseitige Lage haben, d. h. wenn es eine Transformation der Gruppe gibt, die auf die Koordinaten von Σ_ξ angewandt, Σ_ξ überführt in Σ_η, und ebenso auf die Koordinaten von $\Sigma_{'\xi}$ angewandt, $\Sigma_{'\xi}$ überführt in $\Sigma_{'\eta}$. Ebenso folgt aus der Definition der $(-)$-Äquipollenz, daß die zwei Gitterpaare $(-)$-äquipollent sind, wenn sie gleiche *absolute* gegenseitige Lage haben, d. h. wenn es eine Transformation der Gruppe gibt, die, auf die zu Σ_0 gehörigen Koordinaten angewandt, gleichzeitig Σ_ξ in Σ_η und $\Sigma_{'\xi}$ in $\Sigma_{'\eta}$ überführt. Unmittelbar evident ist jetzt, daß die Gitterpaare $(\Sigma_\xi, \Sigma_{'\xi})$ und $(\Sigma_\eta, \Sigma_{'\eta})$ $(\pm)$-äquipollent sind, wenn die Paare $(\Sigma_\xi, \Sigma_\eta)$ und $(\Sigma_{'\xi}, \Sigma_{'\eta})$ $(\mp)$-äquipollent sind.

Bei einer Transformation der *ersten* Parametergruppe $'T = UT$ geht Σ_ξ über in das Gitter, das aus Σ_0 entsteht durch $T_\xi^{-1} U^{-1}$ auf x wirkend, also aus Σ_ξ durch $T_\xi^{-1} U^{-1} T_\xi$ auf x wirkend oder U^{-1} auf $'x$ wirkend. Jedes

Gitter erfährt dabei also in den Koordinaten dieses Gitters selbst die Transformation U^{-1}, alle Gitter erleiden demnach, auf sich selbst bezogen, die gleiche Änderung. Dagegen ändert sich die *absolute* gegenseitige Lage zweier Gitter Σ_ξ und Σ_η nicht, da $T_\eta^{-1} T_\xi$ invariant ist.

Bei einer Transformation der *zweiten* Parametergruppe $'T = TU$ geht Σ_ξ über in das Gitter, das aus Σ_ξ entsteht, wenn U^{-1} auf x wirkt. Daß sich dabei die *relative* gegenseitige Lage zweier Gitter Σ_ξ und Σ_η nicht ändern kann, ist selbstverständlich, folgt aber übrigens auch noch aus der Invarianz von $T_\xi T_\eta^{-1}$.

Wählen wir als Beispiel die sechsgliedrige Gruppe der Bewegungen in R_3. Ein beliebiges rechtwinkliges Koordinatengitter werde mit Σ_0 bezeichnet, Σ_ξ entsteht aus Σ_0 vermöge T_ξ^{-1}. $(\Sigma_\xi, \Sigma_\eta)$ und $(\Sigma_{'\xi}, \Sigma_{'\eta})$ sind $(-)$-äquipollent, wenn es eine Schraubung in R_3 gibt, die gleichzeitig Σ_ξ in $\Sigma_{'\xi}$ und Σ_η in $\Sigma_{'\eta}$ überführt, sie sind $(+)$-äquipollent, wenn $\Sigma_{'\eta}$ aus $\Sigma_{'\xi}$ entstehen kann durch eine Schraubung, die in bezug auf $\Sigma_{'\xi}$ dieselbe Stellung im Raume hat wie eine Σ_ξ in Σ_η überführende Schraubung in bezug auf Σ_ξ. Die geodätische Linie, die die Punkte ξ und η verbindet, bildet alle Gitter ab, die aus Σ_ξ entstehen durch fortgesetzte Ausübung einer infinitesimalen Schraubung, die so gewählt ist, daß schließlich Σ_η entsteht.

Durch die Äquipollenzen werden Segmente einer und derselben geodätischen Linie der Länge nach vergleichbar. Man kann diesen Umstand verwenden zur Definition der „Spiegelung". T_ξ soll das Spiegelbild von $T_{'\xi}$ in bezug auf T_η heißen, wenn $T_\xi^{-1} T_\eta = T_\eta^{-1} T_{'\xi}$. Aus der Definition der Äquipollenzen folgt unmittelbar, daß bei einer Spiegelung die $(\pm)$-Äquipollenz übergeht in die $(\mp)$-Äquipollenz. Ein Punktepaar und sein Spiegelbild sind aber im allgemeinen weder $(+)$- noch $(-)$-äquipollent. Einen besonderen Fall bildet die Spiegelung an $\underset{0}{\xi}$, wobei T_ξ übergeht in T_ξ^{-1}. Wir machen im nächsten Abschnitt gerade von dieser Spiegelung Gebrauch, um von den Eigenschaften der $(+)$-Äquipollenz zu denen der $(-)$-Äquipollenz zu gelangen.

II. Geometrie des Gruppenkeimes.

§ 1.

Die drei Fundamentaltheoreme.

Die drei klassischen Fundamentaltheoreme der Theorie der kontinuierlichen Gruppen lauten:

I. Bilden die Transformationen

$$'\overset{z}{x} = \overset{z}{f}(\overset{y}{x}, \overset{\nu}{\xi}) \qquad (y, z = 1, \ldots, n;\ \nu = a_1, \ldots, a_n) \tag{8}$$

in n Variablen $\overset{z}{x}$ $(z = 1, \dots, n)$ *und r wesentlichen Parametern* $\overset{\nu}{\xi}$ $(\nu = a_1, \dots, a_r)$ *eine Gruppe, so existieren* r^2 *Funktionen* $A^i_\lambda(\xi)$ *der* ξ $(i = 1, \dots, r)$, *deren Determinante* $|A^i_\lambda| \neq 0$ *ist, und* $n\,r$ *Funktionen* $\xi^z_i(x)$ *der* x $(z = 1, \dots, n;\ i = 1, \dots, r)$, *die keiner Gleichung der Form*

$$c^i \xi^z_i = 0 \tag{9}$$

mit konstanten c^i *genügen, so daß*

$$\boxed{\partial_\lambda {}'x^z = \xi^z_j('x)\, A^j_\lambda(\xi)} \tag{10}$$

Ist umgekehrt eine Schar von Transformationen (8) *mit r wesentlichen Parametern gegeben, die die identische Transformation* $\xi = \underset{0}{\xi}$ *enthält und einem System von Gleichungen der Form* (10) *genügt, wo* $|A^z_\lambda(\underset{0}{\xi})| \neq 0$, *so ist die Schar eine Gruppe.*

II. *Sind* X_i *die Symbole von r unabhängigen infinitesimalen Transformationen der Gruppe:*

$$X_i = \xi^\nu_i(\xi)\, \frac{\partial}{\partial x^\nu}, \tag{11}$$

so gelten für die Klammersymbole $(X_i X_j) = X_i X_j - X_j X_i$ *die Gleichungen*

$$\boxed{(X_i X_j) = c_{ij}^{\cdot\cdot k} X_k}, \tag{12}$$

wo die $c_{ij}^{\cdot\cdot k}$ *Konstanten sind, die der Gleichung* $c_{(ij)}^{\cdot\cdot k} = 0$ *genügen. Sind umgekehrt* X_i *die Symbole von r unabhängigen infinitesimalen Transformationen, die den Gleichungen* (12) *genügen, so bilden die Transformationen der durch die* ∞^{r-1} *infinitesimalen Transformationen* $e^i X_i$ *erzeugten eingliedrigen Gruppen eine r-gliedrige Gruppe.*

III. *Die Konstanten* $c_{ij}^{\cdot\cdot k}$ *genügen den Gleichungen*

$$\boxed{c_{[ij}^{\cdot\cdot l}\, c_{k]l}^{\cdot\cdot m} = 0}. \tag{13}$$

Gibt es umgekehrt r^3 *Konstanten* $c_{ij}^{\cdot\cdot k}$, *die den Gleichungen* $c_{(ij)}^{\cdot\cdot k} = 0$ *und* (13) *genügen, so gibt es stets eine r-gliedrige Gruppe, in der man r infinitesimale Transformationen so wählen kann, daß* (12) *gilt.*

§ 2.

Die (+)- bzw. (—)-Äquipollenz der Systeme $\binom{k}{i}$ und $\binom{K}{I}$.

Schreiben wir (10) in der Form

$$d'x^z = \xi^z_j('x)\, A^j_\lambda(\xi)\, d\xi^\lambda, \tag{14}$$

so stellt diese Gleichung die unendlich kleine auf $'x$ angewandte Transformation dar, die nach T_ξ auf x angewandt äquivalent ist mit $T_{\xi+d\xi}$ auf x angewandt, d. i. also $T_{\xi+d\xi}T_\xi^{-1}$. Nun ist $|A_\lambda^k| \neq 0$, und die

$$(15) \qquad (d\xi)^k = A_\lambda^k(\xi)\, d\xi^\lambda \ {}^{5)}$$

sind also Bestimmungszahlen des Linienelementes der X_r in bezug auf ein System $\binom{k}{i}$ von Maßvektoren $\underset{i}{e}$, $\overset{k}{e}$, für deren Bestimmungszahlen in bezug auf das zu den Urvariablen $\overset{\nu}{\xi}$ gehörige System $\binom{\nu}{\lambda}$ folgende skalare Gleichungen gelten:

$$(16) \qquad \overset{k}{e}_\lambda \overset{*}{=} A_\lambda^k;\quad \underset{i}{e}^\nu \overset{*}{=} A_i^\nu;\quad A_i^\alpha A_\alpha^k \overset{*}{=} \begin{cases} 1, & i = k, \\ 0, & i \neq k. \end{cases} \ {}^{5a)}$$

Die A_λ^k sowie die A_i^ν sind Bestimmungszahlen des Einheitsaffinors in bezug auf diese beiden Systeme. Für die Transformation $T_{\xi+d\xi}T_\xi^{-1}$, angewandt auf x, können wir jetzt schreiben:

$$(17) \qquad dx^s = \xi_j^s(x)\,(d\xi)^j.$$

Sind jetzt η^k beliebige andere Werte der Parameter und werden die $d\eta^k$ so gewählt, daß

$$(18) \qquad T_{\eta+d\eta}T_\eta^{-1} = T_{\xi+d\xi}T_\xi^{-1},$$

d. h. daß die beiden Strecken T_η, $T_{\eta+d\eta}$ und T_ξ, $T_{\xi+d\xi}$ $(+)$-äquipollent sind, so folgt aus (17)

$$(19) \qquad (d\xi)^k = (d\eta)^k.$$

$(+)$-äquipollente Strecken haben also in bezug auf das $\binom{k}{i}$-System dieselben Bestimmungszahlen, mit anderen Worten: *das aus dem ersten Teil des ersten Fundamentalsatzes hervorgehende* $\binom{k}{i}$*-System ist in je zwei Punkten* $(+)$-äquipollent [6]) [7]).

---

[5]) Wir schreiben $(d\xi)^k$ statt $d\xi^k$, da die ξ^k für sich im allgemeinen keinen Sinn haben.

[5a]) Wir verwenden das Zeichen $\overset{*}{=}$, um anzudeuten, daß die Ausdrücke links und rechts sich nicht stets in derselben Weise transformieren. Eine Gleichung mit $\overset{*}{=}$ bleibt also nicht bei allen Transformationen richtig und darf z. B. nicht kovariant differenziert werden.

[6]) C & S I, S. 805.

[7]) Es sei kurz daran erinnert, wie sich die ξ_i^s und die A_λ^s aus den Gleichungen der Gruppe berechnen lassen. In den Gleichungen

$$'\overset{s}{x} = \overset{s}{f}(x, \xi),$$

$$''\overset{s}{x} = \overset{s}{f}('x, \eta) = \overset{s}{f}(x, \varphi)$$

(Fortsetzung der Fußnote 7) auf nächster Seite.)

Da auch die Umkehrung von (8)

$$\overset{z}{x} = \overset{z}{F}('\overset{y}{x}, \overset{v}{\xi}) \tag{20}$$

eine Gruppe darstellt, gilt als Gegenstück zu (14)

$$dx^z = \Xi^z_J(x)\, A^J_\lambda(\xi)\, d\xi^\lambda \qquad (J = \bar{1}, \ldots, \bar{r}), \tag{21}$$

die Gleichung der infinitesimalen Transformation $T^{-1}_{\xi+d\xi}\, T_\xi$ in den x. Da $|A^K_\lambda| \neq 0$ ist, sind die

$$(d\xi)^K = A^K_\lambda(\xi)\, d\xi^\lambda \tag{22}$$

Bestimmungszahlen des Linienelementes der X_r in bezug auf ein System $\binom{K}{J}$ von Maßvektoren $\underset{J}{e}$, $\overset{K}{e}$, für deren Bestimmungszahlen die skalaren Gleichungen

$$\overset{K}{e}_\lambda \overset{*}{=} A^k_\lambda;\; \underset{I}{e}^v \overset{*}{=} A^v_I;\; A^\alpha_I A^K_\alpha \overset{*}{=} \begin{cases} 1, & I = K, \\ 0, & I \neq K \end{cases} \tag{23}$$

gelten. Wie oben wird gezeigt, daß das System $\binom{K}{I}$ *in je zwei Punkten* (—)-*äquipollent ist*. Den synthetischen Betrachtungen im ersten Abschnitt zufolge ist dies übrigens unmittelbar evident, da ja bei Spiegelung an $\underset{0}{\xi}$ (+)- und (—)-Äquipollenz vertauscht werden.

Zusammenfassend lesen wir jetzt die Hauptgleichung (14) des ersten Fundamentalsatzes sowie (21) folgendermaßen:

Jedem System von (+)- *bzw.* (—)-*äquipollenten infinitesimalen Strecken der X_r ist in eineindeutiger Weise eine infinitesimale Transformation zugeordnet, eben die Transformation, die mit dem Endpunkte der in $\underset{0}{\xi}$ anfangenden Strecke korrespondiert.*

Die erste der beiden Zuordnungen tritt am schärfsten hervor, wenn wir wie üblich setzen:

$$\xi^y_i(x)\, \frac{\partial f}{\partial x^y} = X_i f, \tag{24}$$

$$A^k_\lambda\, d\xi^\lambda = (d\xi)^k = e^k\, dt; \tag{25}$$

---

betrachtet man x, ξ und φ als unabhängige Variable und differenziert nach den ξ. Dann läßt sich $\partial_\lambda 'x^k$ lösen in der Form

$$\partial_\lambda 'x^z = \Phi^z_j('x, \eta)\, \Psi^j_\lambda(\xi, \eta).$$

Erteilt man in dieser Gleichung den η beliebige, aber feste Werte, so entsteht eine Gleichung der Form (10). Die ξ^k_i und die A^k_λ und somit auch das System $\binom{k}{i}$ lassen sich also rein durch Differentiationen und Eliminationen aus den Gleichungen der Gruppen bestimmen. Verschiedene Wertsysteme der η führen zu verschiedenen Systemen $\binom{k}{i}$.

e^k ist sodann ein $(+)$-konstantes kontravariantes Vektorfeld und $e^k X_k$ das Symbol der zu e^k gehörigen infinitesimalen Transformation:

$$df = e^j X_j f dt. \tag{26}$$

Jede infinitesimale Transformation setzt sich linear mit konstanten Koeffizienten aus den X_j zusammen, und die X_j transformieren sich bei linearen Transformationen des Systems $\binom{k}{i}$ wie die Bestimmungszahlen eines kovarianten Vektors.

Die Größen A_i^ν und A_I^ν stehen in enger Beziehung zu den beiden Parametergruppen. Setzen wir

$$A_i^\mu \frac{\partial}{\partial \xi^\mu} = A_i; \qquad A_I^\mu \frac{\partial}{\partial \xi^\mu} = A_I, \tag{27}$$

so ist A_i das Symbol einer infinitesimalen Transformation in den ξ, die auf ξ^ν angewandt liefert

$$d\xi^\nu = A_i \xi^\nu dt = \underset{i}{e}^\nu dt \tag{28}$$

und also eine Verschiebung darstellt über die in je zwei Punkten $(+)$-äquipollente Strecke $\underset{i}{e}^\nu dt$. Ist U die zu diesem Streckenfeld gehörige infinitesimale Transformation in x, so stellt die Gleichung $'T = UT$ also gerade die infinitesimale Transformation A_i dar. A_i bzw. A_I ist also *Symbol einer infinitesimalen Transformation der ersten bzw. zweiten Parametergruppe* (vgl. S. 247).

§ 3.

Die Größen C und $\overline{C}$ und ihre geometrische Bedeutung.

Sind $\bar{\xi}$ die Parameter der Umkehrung von T_ξ:

$$x^s = \overset{s}{F}('x, \xi) = \overset{s}{f}('x, \bar{\xi}), \tag{29}$$

so ist

$$T^{-1}_{\xi + d\xi} T_\xi = T_{\bar{\xi} + d\bar{\xi}} T_{\bar{\xi}}^{-1}, \tag{30}$$

so daß (21) äquivalent ist mit

$$dx^s = \xi_j^s(x) A_\lambda^j(\bar{\xi}) d\bar{\xi}^\lambda. \tag{31}$$

Für $\xi = \underset{0}{\xi}$ ist aber $d\bar{\xi}^\lambda = -d\xi^\lambda$, und aus (21) und (31) folgt also

$$\Xi_I^s(x) = -\xi_j^s(x) A_I^j(\underset{0}{\xi}). \tag{32}$$

Wir wollen nun die $\binom{K}{I}$-Systeme so wählen, daß in $\underset{0}{\xi}$ das $\binom{K}{I}$-System sich mit dem $\binom{k}{i}$-System deckt. Dann gelten die skalaren Gleichungen

$$A_I^k(\xi_0) \overset{*}{=} \begin{cases} 1, & I = \bar{k}, \\ 0, & I \neq k, \end{cases} \tag{33}$$

und es wird $\Xi_I^\varkappa = -\xi_i^\varkappa$ für $I = \bar{i}$. Ist also $\bar{C}$ eine Größe, deren $\binom{k}{I}$-Bestimmungszahlen in *allen* Punkten gleich $A_I^k(\underset{0}{\xi})$ sind, und C ihre Umkehrung, so daß die skalaren Gleichungen gelten

$$\bar{C}_I^{\cdot k} \overset{*}{=} \begin{cases} 1, & I = \bar{k} \\ 0, & I \neq \bar{k} \end{cases}; \qquad C_i^{\cdot K} \overset{*}{=} \begin{cases} 1, & K = \bar{i} \\ 0, & K \neq \bar{i} \end{cases}, \tag{34}$$

so geht (32) über in

$$\boxed{\Xi_I^\varkappa(x) = -\bar{C}_I^{\cdot j}\,\xi_j^\varkappa(x)} \tag{35}$$

Da infolge (34) z. B. $C_j^{\cdot K} v^j$ einen Vektor darstellt, der in bezug auf $\binom{K}{I}$ dieselben Bestimmungszahlen hat wie v in bezug auf $\binom{k}{i}$, ist die geometrische Bedeutung von C und $\bar{C}$ klar, in jedem Punkte der X_r bewirken das lokale C und $\bar{C}$ die linearen Transformationen, die die Systeme $\binom{k}{i}$ und $\binom{K}{I}$ ineinander überführen.

Die Größen C und $\bar{C}$ stehen, wie wir zeigen wollen, in enger Beziehung zur adjungierten Gruppe. Die infinitesimale Transformation $T_{\xi+d\xi} T_\xi^{-1}$, angewandt auf die $'x$, ist gegeben durch (14). Ferner ist $T_{\xi+d\xi}^{-1} T_\xi$, angewandt auf die x, gegeben durch (21), und die Umkehrung $T_\xi^{-1} T_{\xi+d\xi}$, ebenfalls angewandt auf die x, läßt sich also unter Berücksichtigung von (35) schreiben:

$$dx^\varkappa = \xi_j^\varkappa(x)\,\bar{C}_J^{\cdot j} A_\lambda^J(\xi)\,d\xi^\lambda. \tag{36}$$

Nun ist aber, wie wir auf S. 248 sahen, $T_\xi^{-1} T_{\xi+d\xi}$, angewandt auf die x, äquivalent mit $T_{\xi+d\xi}^{-1} T_\xi$, angewandt auf die $'x$. Ist also f eine Funktion der x und somit auch der $'x$, so ist bei dieser Transformation

$$df = \frac{\partial f}{\partial x^\varkappa}\,\xi_j^\varkappa(x)\,\bar{C}_J^{\cdot j} A_\lambda^J\,d\xi^\lambda = \frac{\partial f}{\partial' x^\varkappa}\,\xi_i^\varkappa('x)\,A_\lambda^i\,d\xi^\lambda, \tag{37}$$

woraus folgt

$$\xi_i^k('x)\,\frac{\partial}{\partial' x^k} = \bar{C}_i^{\cdot j}\,\xi_j^k(x)\,\frac{\partial}{\partial x^k} \tag{38}$$

oder abgekürzt

$$'X_i = \bar{C}_i^{\cdot j}(\xi)\,X_j. \tag{39}$$

Ist nun $e^i X_i$ eine gegebene infinitesimale Transformation in x, und $e^{i\prime} X_i'$ dieselbe Transformation in $'x$ geschrieben, so daß also

$$(40)\qquad e^i X_i f = {}'e^{i\prime} X_i f,$$

so folgt aus (39)

$$(41)\qquad {}'e^k = C_j^{\ k}(\xi)\, e^j$$

oder

$$(42)\qquad {}'e^K(\xi) = C_j^{\ K} e^j .$$

Aus dieser letzten Gleichung folgt, daß $'e^k$ das $(+)$-konstante Vektorfeld ist, dessen $\binom{K}{I}$-Bestimmungszahlen im Punkte ξ den $\binom{k}{i}$-Bestimmungszahlen von e^k gleich sind. Mit anderen Worten:

e^k sei ein $(+)$-konstantes Vektorfeld. Um die infinitesimale Transformation in $'x = T_\xi x$ zu finden, die der zu $e^k dt$ gehörigen infinitesimalen Transformation in x gleichwertig ist, hat man den Vektor e^k in $\underset{0}{\xi}$ $(-)$-äquipollent nach ξ zu verschieben und aus dem so erhaltenen Vektor $'e^k$ ein $(+)$-konstantes Feld zu bilden.

In der Tat, ist T die zu $e^k dt$ gehörige infinitesimale Transformation, so geht die zu $e^k dt$ in $\underset{0}{\xi}$ gehörige Strecke (I, S) bei $(-)$-äquipollenter Verschiebung über in $(T_\xi, T_\xi T)$, und wenn man diese Strecke $(+)$-äquipollent zurück verschiebt, so entsteht $(I, T_\xi T T_\xi^{-1})$.

$T_\xi T T_\xi^{-1}$ ist eine Transformation der adjungierten Gruppe. Bei dieser Gruppe transformieren sich also die infinitesimalen Transformationen linear, und es gilt der Satz:

In ihrer Wirkung auf die infinitesimalen Transformationen ist die adjungierte Gruppe die Gruppe aller durch die Größen $C_i^{\ k}$ in den verschiedenen Punkten von X_r vermittelten Transformationen. Mit T_ξ korrespondiert die durch $C_i^{\ k}$ im Punkte ξ vermittelte Transformation, die das lokale $\binom{k}{i}$-System in das $\binom{K}{I}$-System überführt.

§ 4.

Die Größe c und die drei Übertragungen.

Aus der Transformation der X_i bei Änderung des $\binom{k}{i}$-Systems folgt, daß die im zweiten Fundamentalsatz auftretenden Koeffizienten $c_{ij}^{\ \ k}$ die $\binom{k}{i}$-Bestimmungszahlen einer Größe dritten Grades sind. Da sie konstant

sind, ist diese Größe $(+)$-konstant. Geht man von der Umkehrung (20) aus, so folgen die Gleichungen

$$\Xi_I^z \frac{\partial}{\partial x^z} \Xi_J^y - \Xi_J^z \frac{\partial}{\partial x^z} \Xi_I^y = \bar{c}_{IJ}{}^K \Xi_K^y, \tag{43}$$

wo die $\bar{c}_{IJ}{}^K$ konstant sind. Neben c existiert also noch eine $(-)$-konstante Größe $\bar{c}$ mit den $\binom{K}{I}$-Bestimmungszahlen $\bar{c}_{IJ}{}^K$. Aus (38) folgt direkt, daß die $\binom{K}{I}$-Bestimmungszahlen von $-\bar{c}$ den korrespondierenden $\binom{k}{i}$-Bestimmungszahlen von c gleich sind. Daß aber c und $-\bar{c}$ *überhaupt identisch* sind, daß es also eine Größe c gibt, die in jedem Punkte der X_r dieselben Bestimmungszahlen in bezug auf $\binom{k}{i}$ und auf $\binom{K}{I}$ hat, welche überdies noch in allen Punkten der X_r dieselben sind, wird das wichtigste Resultat dieses Paragraphen sein.

Bei dem bekannten Beweise des ersten Teiles des zweiten Fundamentalsatzes geht man aus von der Tatsache, daß das System der Gleichungen (10) die Lösungen (8) besitzt, wo die x die Rolle von n beliebigen Integrationskonstanten spielen. Das System (10) ist also *vollständig*, und das gleiche gilt infolge eines bekannten Satzes aus der Theorie der Differentialgleichungen für das *adjungierte* System in $n+r$ Variablen

$$\frac{\partial}{\partial \xi^\lambda} \Omega(x', \xi) + \frac{\partial}{\partial' x^z} \Omega('x, \xi)\, \xi_i^z('x)\, A_\lambda^i(\xi) = 0, \tag{44}$$

das sich auch in der Form (vgl. (38), (39) und (27))

$$A_i \Omega + X_i' \Omega = 0 \tag{45}$$

schreiben läßt. Die Integrabilitätsbedingungen von (45) lauten dann

$$(X_i' X_j') = c_{ij}{}^k X_k', \tag{46}$$

$$(A_i A_j) = c_{ij}{}^k A_k, \tag{47}$$

wo die $c_{ij}{}^k$ weder Funktionen der x noch der ξ sein können. Wir erinnern nur an diesen Gedankengang, weil (47) uns unter Berücksichtigung von (27) unmittelbar die $c_{ij}{}^k$ liefert, ausgedrückt in den schon auf S. 251 berechneten A_i^ν und A_λ^k:

$$\boxed{c_{ij}{}^k = -2 A_{[i}^\mu A_{j]}^\lambda \partial_\mu A_\lambda^k} \tag{48}$$

In derselben Weise berechnet man

$$\boxed{\bar{c}_{IJ}{}^K = -2 A_{[I}^\mu A_{J]}^\lambda \partial_\mu A_\lambda^K} \tag{49}$$

Wir berechnen nun zunächst die Parameter $\overset{+}{\Gamma}{}^{\nu}_{\lambda\mu}$ und $\overset{-}{\Gamma}{}^{\nu}_{\lambda\mu}$ der (+)- bzw. (−)-Übertragung. Da A Einheitsaffinor ist, ist

$$0 = \overset{+}{\nabla}_\mu A^\nu_i = \partial_\mu A^\nu_i + \overset{+}{\Gamma}{}^{\nu}_{\lambda\mu} A^\lambda_i, \tag{50}$$

so daß

$$\boxed{\overset{+}{\Gamma}{}^{\nu}_{\lambda\mu} = A^\nu_i \partial_\mu A^i_\lambda} \tag{51}$$

und unter Berücksichtigung von (48)

$$\overset{+}{\Gamma}{}^{\nu}_{[\lambda\mu]} = \frac{1}{2} c_{\lambda\mu}^{\cdot\cdot\nu}. \tag{52}$$

In derselben Weise läßt sich zeigen, daß

$$\boxed{\overset{-}{\Gamma}{}^{\nu}_{\lambda\mu} = A^\nu_I \partial_\mu A^I_\lambda} \tag{53}$$

und

$$\overset{-}{\Gamma}{}^{\nu}_{[\lambda\mu]} = \frac{1}{2} \bar{c}_{\lambda\mu}^{\cdot\cdot\nu}. \tag{54}$$

Es seien jetzt e^ν und s^ν zwei beliebige (+)-konstante Vektorfelder. Wir wollen den Vektor berechnen, der entsteht, wenn e^ν in $\underset{0}{\xi}$ (−)-äquipollent verschoben wird über $s^\nu dt$. Sind S und T die zu $s^\nu dt$ und $e^\nu dt$ gehörigen infinitesimalen Transformationen, so korrespondiert der gesuchte Vektor also mit STS^{-1}. Nun ist

$$Sx^s = x^s + s^j X_j x^s dt + \frac{1}{2} s^i s^j X_i X_j x^s dt^2 + \ldots \tag{55}$$

STx^k (nicht TSx^k! Vgl. die Bemerkung am Schluß des § I 1, S. 245) wird gebildet, indem in diesem Ausdruck T auf x^k angewandt wird und $STS^{-1}x^k$ entsteht ebenso, indem in dem zuletzt erhaltenen Ausdruck S^{-1} auf x^k angewandt wird. Schließlich ergibt sich

$$\begin{aligned} STS^{-1}x^s = x^s + e^j X_j x^s dt + \frac{1}{2} e^i e^j X_i X_j x^s dt^2 + \\ + e^i s^j (X_i X_j - X_j X_i) x^s dt^2 + \ldots, \end{aligned} \tag{56}$$

und diese infinitesimale Transformation stimmt also bis auf Größen dritter und höherer Ordnung überein mit der zum Vektor

$$e^k + e^i s^j c_{ij}^{\cdot\cdot k} dt \tag{57}$$

gehörigen. Der Vektor e^ν in $\underset{0}{\xi}$ geht also bei (−)-äquipollenter Verschiebung über $s^\nu dt$ in $e^\nu + e^\lambda s^\mu c_{\lambda\mu}^{\cdot\cdot\nu} dt$ über. Da derselbe Vektor bei (+)-äquipollenter Verschiebung in e^ν übergeht, ist

$$0 = \overset{+}{\delta} e^\nu = \overset{-}{\delta} e^\nu + c_{\lambda\mu}^{\cdot\cdot\nu} e^\lambda s^\mu dt, \tag{58}$$

woraus folgt, da e^ν und s^ν beliebig gewählt sind,

$$\overset{+}{\Gamma}{}^\nu_{\lambda\mu} = \overset{-}{\Gamma}{}^\nu_{\lambda\mu} + c_{\lambda\mu}^{\cdot\cdot\nu} \tag{59}$$

und unter Berücksichtigung von (52) und (54)

$$c_{\lambda\mu}^{\cdot\cdot\nu} = -\bar{c}_{\lambda\mu}^{\cdot\cdot\nu} = 2\overset{+}{\Gamma}{}^\nu_{[\lambda\mu]} = -2\overset{-}{\Gamma}{}^\nu_{[\lambda\mu]}. \tag{60}$$

Außerdem ergibt sich, daß zu beiden Übertragungen dieselbe *symmetrische* Übertragung gehört, die wir die 0-Übertragung nennen wollen und deren Parameter mit $\overset{0}{\Gamma}{}^\nu_{\lambda\mu}$ bezeichnet werden mögen:

$$\overset{0}{\Gamma}{}^\nu_{\lambda\mu} = \overset{+}{\Gamma}{}^\nu_{(\lambda\mu)} = \overset{-}{\Gamma}{}^\nu_{(\lambda\mu)}. \tag{61}$$

Ein bei dieser Übertragung invariantes Feld soll (0)-*konstant* heißen.

Eine unmittelbare Folge der Identität von c und $\bar{c}$ ist, daß c sowohl (+)- als (−)-konstant ist:

$$\boxed{\overset{+}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} = \overset{-}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} = 0}. \tag{62}$$

Außerdem folgt aus der Gleichheit der Bestimmungszahlen in bezug auf alle Systeme $\binom{k}{i}$ und $\binom{K}{I}$, *daß die* $\binom{k}{i}$*-Bestimmungszahlen von c bei allen Transformationen der adjungierten Gruppe invariant sind.*

Die drei Übertragungen lassen sich natürlich auch in bezug auf die Systeme $\binom{k}{i}$ bzw. $\binom{K}{I}$ festlegen[8]). Nennt man die zugehörigen Parameter λ^k_{ij} bzw. Λ^K_{IJ} mit den oberen Indizes 0, + und −, so berechnet sich leicht

$$\begin{array}{ll} \overset{+}{\lambda}{}^k_{ij} = 0 & \overset{+}{\Lambda}{}^K_{IJ} = c_{IJ}^{\cdot\cdot K} \\ \overset{0}{\lambda}{}^k_{ij} = -\frac{1}{2} c_{ij}^{\cdot\cdot k} & \overset{0}{\Lambda}{}^K_{IJ} = \frac{1}{2} c_{IJ}^{\cdot\cdot K} \\ \overset{-}{\lambda}{}^k_{ij} = -c_{ij}^{\cdot\cdot k} & \overset{-}{\Lambda}{}^K_{IJ} = 0, \end{array} \tag{63}$$

so daß z. B.

$$\begin{aligned} \overset{+}{\delta} v^k &= dv^k \\ \overset{0}{\delta} v^k &= dv^k - \frac{1}{2} c_{ij}^{\cdot\cdot k} v^i (d\xi)^j \\ \overset{-}{\delta} v^k &= dv^k - c_{ij}^{\cdot\cdot k} v^i (d\xi)^j. \end{aligned} \tag{64}$$

---

[8]) Für die Festlegung von Übertragungen mit Hilfe von nichtholonomen Bezugssystemen vergleiche man die Arbeit des Verfassers Über nichtholonome Übertragungen in einer L_n, Math. Zeitschr. **30** (1928), S. 1—24, § 2.

§ 5.

Kanonische Parameter und Reihenentwicklungen.

Die endlichen Transformationen der durch die infinitesimale Transformation $e^i X_i$ erzeugten eingliedrigen Untergruppe lauten bekanntlich

$$(65)\qquad {}'x^s = x^s + \frac{1}{1!}\, e^i X_i x^s t + \frac{1!}{2!}\, e^i e^j X_i X_j x^s t^2 + \ldots$$

oder, wenn wir $e^i t$ durch ε^i ersetzen,

$$(66)\qquad {}'x^s = x^s + \frac{1}{1!}\, \varepsilon^i X_i x^s + \frac{1}{2!}\, \varepsilon^i \varepsilon^j X_i X_j x^s + \ldots .$$

Bei veränderlichen ε stellen diese Gleichungen *sämtliche* Transformationen des Gruppenkeims dar, die ε lassen sich also als neue Parameter, sogenannte *kanonische Parameter*, einführen. *Da im Ricci-Kalkül aber die Art der Indizes das Bezugssystem angibt* und z. B. die v^k nun einmal die Bestimmungszahlen in bezug auf das System $\binom{k}{i}$ sind, muß für die kanonischen Parameter eine neue Art von Indizes eingeführt werden. Wir verwenden *gothische* Indizes und schreiben für das zu den kanonischen Parametern gehörige System $\binom{\mathfrak{k}}{\mathfrak{i}}$, das nur in $\underset{0}{\xi}$ mit dem System $\binom{k}{i}$ zusammenfällt, $\underset{\mathfrak{i}}{e}^{\mathfrak{k}}$, $\overset{\mathfrak{k}}{e}_{\mathfrak{i}}$, $\mathfrak{i}, \mathfrak{j}, \mathfrak{k} = 1', \ldots, n'$. Nur für die ε selbst und ihre Differentiale soll *neben* $\varepsilon^{\mathfrak{k}}$ und $d\varepsilon^{\mathfrak{k}}$ auch ε^k und $d\varepsilon^k$ zugelassen bleiben, da die Bestimmungszahlen des Linienelements in bezug auf $\binom{k}{i}$ doch $(d\varepsilon)^k$ geschrieben werden und Verwirrung also ausgeschlossen ist.

Da

$$(67)\qquad \underset{\mathfrak{i}}{e}^{k} = \underset{\mathfrak{i}}{e}^{\mathfrak{j}} A_{\mathfrak{j}}^{k};\quad \overset{\mathfrak{k}}{e}_{i} = \overset{\mathfrak{k}}{e}_{\mathfrak{j}} A_{i}^{\mathfrak{j}},$$

sind die Systeme $\binom{\mathfrak{k}}{\mathfrak{i}}$ bekannt, wenn es gelingt, die $A_{\mathfrak{i}}^{k}$ und $A_{i}^{\mathfrak{k}}$ als Funktionen der ξ zu bestimmen. Dies geschieht am einfachsten mit Hilfe der von F. Schur gegebenen Reihenentwicklungen[9]). Wir verwenden die Gleichung (48) in einer Form, die zuerst von Maurer angegeben wurde,

$$(68)\qquad \partial_\mu A_\lambda^k - \partial_\lambda A_\mu^k = c_{\lambda\mu}^{\cdot\cdot k}\,{}^{10)}.$$

Die Gleichung

$$(69)\qquad \varepsilon^k = e^k t$$

---

[9]) Math. Ann. **38** (1891), S. 263—286.

[10]) Münch. Ber. **18** (1888), S. 117.

stellt bei konstanten e^k eine Kurve durch $\underset{0}{\xi}$ dar, die das Abbild der durch $e^k dt$ erzeugten eingliedrigen Gruppe ist. Differentiation ergibt

$$d\varepsilon^k = e^k dt, \tag{70}$$

woraus u. a. hervorgeht, daß die Kurve eine geodätische Linie durch $\underset{0}{\xi}$ ist. Dieser Linie entlang bleibt also nicht nur die $\binom{\mathfrak{l}}{\mathfrak{i}}$-Bestimmungszahl, sondern auch die $\binom{k}{i}$-Bestimmungszahl des Linienelementes konstant:

$$e^{\mathfrak{l}} dt = \varepsilon_j^{\,\mathfrak{l}} e^j dt, \text{ [10a])} \tag{71}$$

so daß

$$\varepsilon_j^{\,\mathfrak{l}} A_{\mathfrak{l}}^{k} e^j dt = e^k dt; \tag{72}$$

oder

$$A_{\mathfrak{l}}^{k} \varepsilon^{\mathfrak{l}} = \varepsilon_{\mathfrak{l}}^{\,k} \varepsilon^{\mathfrak{l}}. \tag{73}$$

Differentiation der letzten Gleichung ergibt, unter Berücksichtigung der Maurerschen Gleichung (68), die für beliebige Parameter ξ, also auch für die kanonischen Parameter ε gilt,

$$\boxed{\frac{dA_{\mathfrak{l}}^{k}}{dt} + \frac{1}{t}(A_{\mathfrak{l}}^{k} - \varepsilon_{\mathfrak{l}}^{\,k}) = - E_{j}^{\cdot k} A_{\mathfrak{l}}^{j}}, \tag{74}$$

wo abkürzend gesetzt ist

$$\boxed{E_{i}^{\cdot k} = - c_{ij}^{\cdot\cdot k} e^j}. \tag{75}$$

(74) ist der gewöhnlichen Differentialgleichung

$$\frac{dx}{dt} + \frac{1}{t}(x - \underset{0}{x}) = \alpha x \tag{76}$$

mit der Lösung

$$x = \underset{0}{x}\left(1 + \frac{1}{2!}\alpha t + \frac{1}{3!}\alpha^2 t^2 + \ldots\right) = \underset{0}{x}\frac{e^{\alpha t} - 1}{\alpha t} \tag{77}$$

analog. Ihre Lösung wird gegeben durch die analoge Reihe

$$A_{\mathfrak{l}}^{k} = \varepsilon_{\mathfrak{l}}^{\,j}\left(A_j^{k} - \frac{1}{2!}E_j^{\cdot k} + \frac{1}{3!}E_j^{\cdot l} E_l^{\cdot k} t^2 - \ldots\right), \tag{78}$$

deren Konvergenz für den Fall, daß die absoluten Werte der $c_{ij}^{\cdot\cdot k}$ unterhalb einer festen Grenze liegen, bewiesen werden kann. In derselben Weise

---

[10a]) Das erweiterte Kroneckersche Symbol ε bedeutet 1 oder 0, je nachdem die beiden Indizes ein jeder in seiner eigenen Zeichenreihe korrespondierende Stellen einnehmen oder nicht.

ergibt sich für $A_i^{\mathfrak{i}}$

$$A_i^{\mathfrak{i}} = \varepsilon_j^{\mathfrak{i}}(A_i^j - \lambda_1 E_i^{\cdot j} t + \lambda_2 E_i^{\cdot l} E_l^{\cdot j} t^2 - \ldots) \tag{79}$$

$$\lambda_q = (-1)^q \frac{B_q}{q!},$$

wo die B_q die Bernoullischen Zahlen sind. Abkürzend schreiben wir

$$\boxed{A_{\mathfrak{i}}^k = \varepsilon_{\mathfrak{i}}^j \left(\frac{e^{-Et} - A}{-Et}\right)_j^k; \quad A_i^{\mathfrak{i}} = \varepsilon_j^{\mathfrak{i}} \left(\frac{-Et}{e^{-Et} - A}\right)_i^j} . \tag{80}$$

Für $A_{\mathfrak{i}}^K$ und $A_I^{\mathfrak{i}}$ ergeben sich ähnliche Ausdrücke, die $\overline{E}_J^{\cdot K} = - \bar{c}_{IJ}^{\cdot\cdot K} e^I$ enthalten. Da aber infolge der Identität von c und $-\bar{c}$ auch E und $-\overline{E}$ identisch sind, folgt

$$\boxed{A_{\mathfrak{i}}^K = \varepsilon_{\mathfrak{i}}^J \left(\frac{e^{Et} - A}{Et}\right)_J^K; \quad A_I^{\mathfrak{i}} = \varepsilon_J^{\mathfrak{i}} \left(\frac{Et}{e^{Et} - A}\right)_I^J} . \tag{81}$$

Aus (80) und (81) läßt sich eine Reihenentwicklung für $C_i^{\cdot k}$ ableiten

$$\begin{aligned} C_i^{\cdot k} &= C_i^{\cdot J} A_J^k = C_i^{\cdot J} A_{\mathfrak{i}}^k A_J^{\mathfrak{i}} = (e^{-Et})_i^k \\ &= A_i^k - E_i^{\cdot k} t + \frac{1}{2!} E_i^{\cdot j} E_j^{\cdot k} t^2 - \ldots, \end{aligned} \tag{82}$$

die sich auf anderem Wege auch aus den Differentialgleichungen der adjungierten Gruppe ergeben wird.

Um die ξ als Funktionen der ε zu berechnen, schreiben wir (70) in der Form

$$\frac{d\xi^\nu}{dt} = e^\nu = A_i^\nu(\xi) e^i \tag{83}$$

und differenzieren fortgesetzt nach t, wobei die e^i Konstanten sind. Für ξ^ν ergibt sich dann die Reihenentwicklung

$$\boxed{\xi^\nu = \underset{0}{\xi^\nu} + A_i^\nu(\underset{0}{\xi}) \varepsilon^i + \frac{1}{2}(A_j^\mu \partial_\mu A_i^\nu)(\underset{0}{\xi}) \varepsilon^i \varepsilon^j + \ldots} . \tag{84}$$

Veblen und Thomas [11]) gaben eine Reihenentwicklung für die Punkte einer geodätischen Linie durch $\underset{0}{\xi}$ in einer allgemeinen A_r, ausgehend von der Differentialgleichung der geodätischen Linien

$$\frac{d^2\xi^\nu}{dt^2} + \overset{0}{\Gamma}{}_{\lambda\mu}^{\nu} \frac{d\xi^\lambda}{dt} \frac{d\xi^\mu}{dt} = 0. \tag{85}$$

Durch fortgesetzte Differentiation dieser Gleichung ergibt sich für eine

---

[11]) The geometry of paths, Trans. Amer. Math. Soc. **25** (1923), p. 551—600, p. 560.

geodätische Linie durch $\underset{0}{\xi}$, deren Richtung in $\underset{0}{\xi}$ durch $\left(\frac{d\xi^\nu}{dt}\right)_0 = \underset{0}{e}^\nu$ festgelegt ist,

$$(86)\quad \boxed{\xi^\nu = \underset{0}{\xi}^\nu + \underset{0}{e}^\nu t - \frac{1}{2!} A^\nu_{\lambda\mu}(\underset{0}{\xi}) \underset{0}{e}^\lambda \underset{0}{e}^\mu t^2 - \frac{1}{3!} A^\nu_{\lambda\mu\omega}(\underset{0}{\xi}) \underset{0}{e}^\lambda \underset{0}{e}^\mu \underset{0}{e}^\omega t^3 - \dots},$$

in welchen Gleichungen die Koeffizienten A durch folgende Rekursionsformel gegeben sind:

$$(87)\quad A^\nu_{\lambda_1 \dots \lambda_{p+1}} = \left(\frac{\partial}{\partial \xi^{\lambda_{p+1}}} A^\nu_{\lambda_1 \dots \lambda_p} - p\, A^\nu_{\alpha(\lambda_1 \dots \lambda_{p-1}} A^\alpha_{\lambda_p \lambda_{p+1})}\right);\quad A^\nu_{\lambda\mu} = \overset{0}{\Gamma}{}^\nu_{\lambda\mu}.$$

Läßt man das $\binom{k}{i}$-System in $\underset{0}{\xi}$ mit dem $\binom{\nu}{\lambda}$-System zusammenfallen, so wird $\underset{0}{e}^\nu t = \varepsilon^\nu$, und mit Hilfe von (61) und (51) läßt sich die Identität der beiden Reihen (84) und (86) Glied für Glied nachweisen. Die kanonischen Parameter sind identisch mit den *Normalkoordinaten* $\underset{0}{e}^\nu t$ von Veblen und Thomas, die für den speziellen Fall, daß die A_r in eine Riemannsche Mannigfaltigkeit übergeht, mit den Riemannschen Normalkoordinaten zusammenfallen.

§ 6.

Verschiedene Deutungen des ersten Teiles des dritten Fundamentaltheorems.

Da die Gleichung (13) die invariante Form hat, gilt sie in bezug auf jedes Bezugssystem, also auch in bezug auf das System $\binom{\nu}{\lambda}$

$$(88)\quad \boxed{c_{[\omega\mu}^{\cdot\cdot\alpha}\, c_{\lambda]\alpha}^{\cdot\cdot\nu} = 0}.$$

In dieser Form leitet sie sich aber unmittelbar ab aus der (+)- und (−)-Konstanz von c (vgl. (62)), denn infolge (59) ist

$$(89)\quad \overset{+}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} = \overset{-}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} + 3\, c_{[\omega\mu}^{\cdot\cdot\alpha}\, c_{\lambda]\alpha}^{\cdot\cdot\nu}.$$

Da aber infolge (59) und (61)

$$(90)\quad \overset{0}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} = \overset{-}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} + \frac{3}{2}\, c_{[\omega\mu}^{\cdot\cdot\alpha}\, c_{\lambda]\alpha}^{\cdot\cdot\nu},$$

folgt, daß auch die (0)-Ableitung von $c_{\lambda\mu}^{\cdot\cdot\nu}$ verschwindet:

$$(91)\quad \boxed{\overset{0}{\nabla}_\omega c_{\lambda\mu}^{\cdot\cdot\nu} = 0}.$$

Die Größe c ist also sowohl (+)- *als* (−)- *als* (0)-*konstant.*

Berechnet man die Krümmungsgrößen der $(+)$- und der $(-)$-Übertragung, so folgt

$$(92)\qquad \overset{+}{R}_{\omega\mu\lambda}{}^{\nu}-\overset{-}{R}_{\omega\mu\lambda}{}^{\nu}=-3\,c_{[\omega\mu}{}^{\alpha}\,c_{\lambda]\alpha}{}^{\nu}=0\,.$$

Da aber beide Übertragungen integrabel sind, verschwindet sowohl $\overset{+}{R}$ als $\overset{-}{R}$ und es ergibt sich also auch aus diesem Grunde (88). Ferner findet man für die Krümmungsgröße der nicht integrablen aber symmetrischen (0)-Übertragung

$$(93)\qquad \boxed{\overset{0}{R}_{\omega\mu\lambda}{}^{\nu}=\frac{1}{4}\,c_{\omega\mu}{}^{\alpha}\,c_{\alpha\lambda}{}^{\nu}}$$

und (88) ist also schließlich auch noch identisch mit der für jede symmetrische Übertragung gültigen sogenannten zweiten Identität $\overset{0}{R}_{[\omega\mu\lambda]}{}^{\nu}=0$. [12])

III. Die adjungierte Gruppe.

§ 1.

Transformation der infinitesimalen und der endlichen Transformationen.

Aus (57) folgt für die mit $s^i X_i$ korrespondierende infinitesimale Transformation der adjungierten Gruppe

$$(94)\qquad d\,e^k=e^i\,s^j\,c_{ij}{}^{k}\,dt.$$

Ist also f eine Funktion der e^k, so ist

$$(95)\qquad df=s^j E_j f,$$

wo (vgl. (75))

$$(96)\qquad E_j=e^i\,c_{ij}{}^{k}\,\frac{\partial}{\partial e^k}=E_j{}^{k}\,\frac{\partial}{\partial e^k}$$

das Symbol einer infinitesimalen Transformation ist. Diese infinitesimale Transformation erzeugt eine eingliedrige Gruppe, deren endliche Transformationen erhalten werden durch Lösung der aus (94) entstehenden Differentialgleichung

$$(97)\qquad \frac{d'e^k}{dt}='e^i\,s^j\,c_{ij}{}^{k}=-'e^i\,S_i{}^{k};\qquad S_j^k=s^i\,c_{ij}{}^{k}$$

mit den Integrationsbedingungen $t=0$, $'e^k=e^k$. Die Lösung von (97) lautet

$$(98)\qquad 'e^k=e^j\,(e^{-St})_j^k=e^j\left(A_j^k-S_j{}^{k}t+\frac{1}{2!}S_j{}^{l}S_l{}^{k}t^2-\ldots\right).$$

Nun ist aber anderseits (vgl. (41))

$$(99)\qquad 'e^k=e^j\,C_j{}^{k}(\xi),$$

[12]) Vgl. R. K., S. 88.

wo für ξ der Wert zu setzen ist, der mit $\varepsilon^i = s^i t$ korrespondiert, und wir haben in (98) also auf anderem Wege die Reihenentwicklung (82) für $C_i^{\cdot k}$ zurückgefunden. Aus (99) folgt ferner

$$'\varepsilon^k = \varepsilon^j C_j^{\cdot k} \tag{100}$$

und C vermittelt also bei Verwendung von ***kanonischen*** Parametern auch die Transformation der ***endlichen*** Transformationen der Gruppe.

§ 2.

Geometrische Deutung in E_r und X_r.

Die kanonischen Parameter ε^k können gedeutet werden als kartesische Koordinaten einer E_r oder als Normalkoordinaten der Gruppenmannigfaltigkeit in bezug auf $\underset{0}{\xi}$ und auf die symmetrische (0)-Übertragung. Bei der ersten Deutung ist die adjungierte Gruppe eine Gruppe von linearen homogenen Transformationen der E_r, die $\underset{0}{\xi}$ invariant lassen. Bekanntlich stellt jede Gerade durch $\underset{0}{\xi}$ eine eingliedrige Untergruppe dar, jede E_m durch $\underset{0}{\xi}$ aber dann und nur dann eine m-gliedrige Untergruppe, wenn sie bei allen Transformationen der adjungierten Gruppe, deren Bildpunkt in dieser E_m liegt, invariant ist, und dann und nur dann eine m-gliedrige invariante Untergruppe, wenn sie bei allen Transformationen der adjungierten Gruppe invariant ist.

Vermöge (96) ist jeder infinitesimalen Transformation der Gruppe eine infinitesimale Transformation der adjungierten Gruppe zugeordnet, oder, was dasselbe ist, vermöge (75) jedem Vektor e in E_r eine Matrix E in E_r. Schreiben wir mit Weyl kleine Buchstaben für die Vektoren, die korrespondierenden großen Buchstaben für die in dieser Weise zugeordneten Matrizen, St für den Vektor, der durch Anwendung von S auf t entsteht und (s, t) für den Vektor $s^i t^j c_{ij}^{\cdot\cdot k}$, so folgt aus (75)

$$St = -Ts = (s, t). \tag{101}$$

Wird ferner (S, T) geschrieben für $ST - TS$, und ist

$$(s, t) = u, \tag{102}$$

so folgt

$$(S, T) = U. \tag{103}$$

Die Gleichung (101) läßt sich folgendermaßen aussprechen:

Der Punkt t der E_r erleidet bei Anwendung der mit dem Punkte s korrespondierenden infinitesimalen Transformation der Gruppe eine infinitesimale Verschiebung $(s, t)\,dt$.

Zur Deutung in X_r sei zunächst bemerkt, daß die eingliedrigen Untergruppen und somit die Geraden durch $\underset{0}{\xi}$ in E_r sich abbilden auf geodätische Linien durch $\underset{0}{\xi}$, daß aber *andere* Geraden in E_r sich im allgemeinen *nicht* auf geodätische Linien der X_r abbilden und daß der gewöhnliche Parallelismus in der E_r nichts mit den beiden Parallelismen in der X_r zu tun hat und sogar überhaupt keine für die Gruppe interessante Bedeutung besitzt. Zur Deutung in X_r betrachten wir die infinitesimalen Transformationen $\underset{a}{t}^k X_k$ $(a, \ldots, g = 1, \ldots, m)$ einer m-gliedrigen Untergruppe. Jedes der m Vektorfelder $\underset{a}{t}^\nu$ ist $(+)$-konstant

$$(104) \qquad \overset{+}{\delta}\, \underset{a}{t}^\nu = 0$$

und bildet eine Kongruenz von geodätischen Linien. Aus (104) und (52) folgt

$$(105) \qquad \underset{a}{t}^\mu \partial_\mu \underset{b}{t}^\nu - \underset{b}{t}^\mu \partial_\mu \underset{a}{t}^\nu = -\underset{a}{t}^\mu c_{\lambda\mu}^{\cdot\cdot\nu} \underset{b}{t}^\lambda,$$

und da die infinitesimalen Transformationen eine Gruppe bilden, liegt der Vektor rechts in dieser Gleichung in der durch die $\underset{a}{t}^\nu$ bestimmten m-Richtung. Die Gleichungen

$$(106) \qquad \underset{a}{t}^\mu \partial_\mu f = 0$$

bilden also ein vollständiges System, oder, was dasselbe ist, die erwähnte m-Richtung ist X_m-bildend. Die Untergruppe bildet sich ab in der X_m durch $\underset{0}{\xi}$.

Jede X_m enthält m Kongruenzen von ∞^{m-1} geodätischen Linien der X_r in den Richtungen der $\underset{a}{t}^\nu$. Da sich aber die infinitesimalen Transformationen der Untergruppe in jeder Richtung der X_m wählen lassen, enthält jede X_m geodätische Linien der X_r durch jeden ihrer Punkte in jeder ihrer Richtungen. Die X_m sind also *geodätisch*. Verschieben wir in irgend einem Punkte P der X_m durch $\underset{0}{\xi}$ einen in P mit $\underset{a}{t}^\nu$ zusammenfallenden Vektor $\overset{*}{\underset{a}{t}}{}^\nu$ $(+)$- bzw. $(-)$-parallel längs $d\xi^\nu = \underset{b}{t}^\nu\, dt$, so ist

$$(107) \qquad d\overset{*}{\underset{a}{t}}{}^\nu = d\underset{a}{t}^\nu \quad \text{bzw.} \quad d\overset{*}{\underset{a}{t}}{}^\nu = d\underset{a}{t}^\nu + c_{\lambda\mu}^{\cdot\cdot\nu}\, \underset{a}{t}^\lambda \underset{b}{t}^\mu\, dt$$

und $\overset{*}{\underset{a}{t}}{}^\nu$ bleibt also bei beiden Verschiebungen innerhalb der X_m.

Ist umgekehrt in X_r eine X_m durch $\underset{0}{\xi}$ gegeben, die den aufgezählten Bedingungen genügt, so kann man in $\underset{0}{\xi}$ m Vektoren $\underset{a}{t}^\nu$ definieren, die die m-Richtung von X_m bestimmen. Bei $(+)$-paralleler Verschiebung entstehen

daraus m $(+)$-konstante Felder $\underset{a}{t}^\nu$, und der Voraussetzung nach bestimmen diese Vektoren in jedem Punkte der X_m die m-Richtung der X_m. Da nun aber die X_m geodätisch ist, liegt $\underset{a}{t}^\mu \overset{0}{\nabla}_\mu \underset{b}{t}^\nu$ stets in einer Richtung der X_m und infolgedessen ist

$$\underset{a}{t}^\mu \overset{0}{\nabla}_\mu \underset{b}{t}^\nu - \underset{b}{t}^\mu \overset{0}{\nabla}_\mu \underset{a}{t}^\nu = - \underset{a}{t}^\mu c_{\lambda\mu}^{\cdot\cdot\nu} \underset{b}{t}^\lambda \tag{108}$$

ein Vektor in einer Richtung der X_m. Die infinitesimalen Transformationen $\underset{a}{t}^k X_k$ bilden also eine Gruppe. Wir haben also den Satz[13]) erhalten:

Die X_m, die erzeugt wird durch die zu einer m-Richtung in $\underset{0}{\xi}$ gehörigen geodätischen Linien der X_r, stellt dann und nur dann eine Untergruppe dar, wenn sie

1. *geodätisch ist und*

2. *jede ihrer Richtungen in jedem Punkte sowohl durch* $(+)$-*parallele als durch* $(-)$-*parallele Verschiebung in einer zur X_m gehörigen Richtung stets in eine zur X_m gehörige Richtung übergeht.*

Es ist zu beachten, daß es nicht hinreicht, daß die X_m geodätisch ist. Denn dadurch wird nur sichergestellt, daß eine Richtung bei (0)-paralleler Verschiebung Richtung der X_m bleibt, und es kann sich demnach der Fall ereignen, daß bei $(+)$- oder $(-)$-paralleler Verschiebung die Richtung aus der X_m heraustritt. Man kann den Unterschied auch folgendermaßen charakterisieren:

Eine geodätische X_m ist in sich (0)-*parallel, stellt sie aber eine Untergruppe dar, so ist sie auch in sich* $(+)$- *und* $(-)$-*parallel.*

Wir wenden uns jetzt zu dem Fall einer invarianten Untergruppe. Um über die Abbildung in X_r Aufschluß zu erhalten, bilden wir m Felder $\underset{a}{\bar t}^\nu$, die in $\underset{0}{\xi}$ mit den $\underset{a}{t}^\nu$ zusammenfallen, aber $(-)$-konstant sind. Sodann ist in $\underset{0}{\xi}$

$$\overset{+}{\delta} \underset{a}{\bar t}^\nu = \overset{-}{\delta} \underset{a}{\bar t}^\nu + c_{\lambda\mu}^{\cdot\cdot\nu} \underset{a}{\bar t}^\lambda d\xi^\mu = c_{\lambda\mu}^{\cdot\cdot\nu} \underset{a}{\bar t}^\lambda d\xi^\mu = c_{\lambda\mu}^{\cdot\cdot\nu} \underset{a}{t}^\lambda d\xi^\mu. \tag{109}$$

Der Fall, daß $d\xi^\nu$ in der X_m liegt, ist oben schon erörtert. Ist $d\xi^\nu$ allgemein gewählt, so folgt aus (109), daß $\overset{+}{\delta} \underset{a}{\bar t}^\nu$ dann und nur dann für jede Wahl von $d\xi^\nu$ in der X_m liegt, wenn die $\underset{a}{t}^k X_k$ eine invariante Untergruppe bilden. *Eine invariante Untergruppe ist also dadurch charakterisiert, daß die zur X_m der Gruppe* $(+)$-*parallelen und* $(-)$-*parallelen X_m sich decken.*

---

[13]) C & S I, S. 813.

Folgende Tabelle erleichtere die Übersicht:

	Lineare Schar von infinitesimalen Transformationen	Untergruppe	Invariante Untergruppe
In E_r:	E_m	E_m, invariant bei allen Transformationen der korrespondierenden Untergruppe der adjungierten Gruppe	E_m invariant bei allen Transformationen der adjungierten Gruppe
In X_r:	In $\underset{0}{\xi}$ geodätische X_m	In sich (+)- und (−)-parallele geodätische X_m	In sich (+)- und (−)-parallele geodätische X_m, zu der jede (+)-parallele X_m auch (−)-parallel ist

§ 3.

Invarianten und Komitanten der adjungierten Gruppe.

Wir erinnern kurz daran, daß die adjungierte Gruppe r-gliedrig ist, wenn der i-Rang[14]) von $c_{ij}^{\cdot\cdot k}$ gleich r ist und daß infolge (96) stets

$$(E_i, E_j) = c_{ij}^{\cdot\cdot k} E_k \tag{110}$$

ist, so daß in diesem Falle die Gruppe und ihre Adjungierte gleichzusammengesetzt sind. Ist dagegen der i-Rang von $c_{ij}^{\cdot\cdot k}$ gleich $q < r$, so gibt es gerade $r - q$ unabhängige *ausgezeichnete* infinitesimale Transformationen, das sind Transformationen, deren Klammerausdrücke mit sämtlichen anderen verschwinden.

Komitante einer Größe soll jede Größe heißen, deren Bestimmungszahlen sich in vom Bezugssystem unabhängiger Weise aus den Bestimmungszahlen der ersten Größe berechnen lassen. Im oben erwähnten Falle $q < r$ gibt es z. B. $r - q$ kontravariante Vektoren, die mit $c_{ij}^{\cdot\cdot k}$ über i überschoben Null erzeugen. Diese bilden eine E_{r-q} und ein kontravarianter $(r-q)$-Vektor in dieser E_{r-q} ist bis auf einen Zahlenfaktor eine Komitante von $c_{ij}^{\cdot\cdot k}$.

Komitante einer Gruppe von linearen homogenen Transformationen heißt jede Größe, deren Bestimmungszahlen sich nicht ändern, wenn das Bezugssystem den Transformationen der Gruppe unterworfen wird. Es ist also sowohl c als auch jede Komitante dieser Größe eine Komitante der

[14]) Gibt es gerade m linear unabhängige Vektoren $\underset{a}{w}^\nu$ $(a = 1, \ldots, m)$, so daß $\underset{a}{w}^{\lambda_1} v_{\lambda_1 \ldots \lambda_p} = 0$ ist, so sagen wir, daß der λ_1-*Rang* von $v_{\lambda_1 \ldots \lambda_p}$ gleich $r - m$ ist. Die Gesamtheit der ∞^{r-m} kovarianten Vektoren, deren Überschiebung mit den w Null ist, bilden das λ_1-*Gebiet* von $v_{\lambda_1 \ldots \lambda_p}$.

adjungierten Gruppe, da ja, wie wir gesehen haben (S. 258), die Bestimmungszahlen $c_{ij}^{\cdot\cdot k}$ bei dieser Gruppe invariant sind. Da aber die adjungierte Gruppe vermöge (96) durch die $c_{ij}^{\cdot\cdot k}$ vollständig festgelegt ist, sind damit *alle* Komitanten erschöpft:

Die Komitanten der adjungierten Gruppe sind die Komitanten der Größe c.

Ferner gilt der Satz:

Wird eine Komitante oder eine Größe, die bis auf einen Zahlenfaktor Komitante der adjungierten Gruppe ist, in beliebiger Weise invariant verknüpft mit beliebig vielen Faktoren e^{ν}, *so entsteht bei Nullsetzen ein bei der adjungierten Gruppe invariantes Gleichungssystem.*

In der Tat, schreiben wir für die erwähnte Verknüpfung symbolisch $\{v, e, e, \ldots\}$, so ist

$$\{v, e, e, \ldots\} = 0, \tag{111}$$

und daraus entsteht bei einer Transformation der adjungierten Gruppe

$$\{v, {}'e, {}'e, \ldots\} = 0, \tag{112}$$

wo der Akzent angeben soll, daß die Bestimmungszahlen in bezug auf $\binom{K}{I}$ einzusetzen sind an Stelle der Bestimmungszahlen in bezug auf $\binom{k}{i}$. Da aber v nach Voraussetzung in bezug auf beide Systeme proportionale Bestimmungszahlen hat, ist (112) gleichwertig mit

$$\{{}'v, {}'e, {}'e, \ldots\} = 0, \tag{113}$$

w. z. b. w.

Ist die Komitante kovariant p-ten Grades und besteht die Verknüpfung aus einer Überschiebung mit p Faktoren e, so entsteht eine *Invariante*. Jede Invariante ist bekanntlich Lösung des Systems

$$e^{i} c_{ij}^{\cdot\cdot k} \frac{\partial f}{\partial e^{k}} = E_{j}^{\cdot k} \frac{\partial f}{\partial e^{k}} = 0, \tag{114}$$

und es gibt also sicher wenigstens eine Invariante, da $e^{j} E_{j}^{\cdot k} = 0$ ist und der Rang von E infolgedessen $\leqq r - 1$ ist.

Ist das entstehende Gleichungssystem linear in e^{k}, so ist dasselbe entweder inkonsistent oder es stellt eine bei der adjungierten Gruppe invariante ebene Mannigfaltigkeit in E_r dar, d. h. also eine *invariante Untergruppe*. Umgekehrt ist ein kontravarianter m-Vektor in der E_m einer invarianten Untergruppe bis auf einen Zahlenfaktor eine Komitante von c.

§ 5.

Die wichtigsten Komitanten und invarianten Untergruppen.

c selbst gibt Anlaß zur Bildung von zwei invarianten Untergruppen. Das Gebiet des Index i bestimmt das *Zentrum*

$$e^i c_{ij}^{\ \ k} = 0, \tag{115}$$

die Gruppe aller infinitesimalen Transformationen, deren Klammerfaktoren mit jeder beliebigen infinitesimalen Transformation der Gruppe verschwinden. Das Zentrum verschwindet nur, wenn der i-Rang von $c_{ij}^{\ \ k}$ gleich r ist, also wenn die Gruppe mit ihrer adjungierten Gruppe gleichzusammengesetzt ist (S. 267).

Das Gebiet des Index k bestimmt die *erste Ableitung*, definiert durch die Gleichung

$$c_{i_1 j_1}^{\cdot\ \cdot\ [k_1} \ldots c_{i_q j_q}^{\cdot\ \cdot\ k_q} e^{k]} = 0 \qquad (q = k\text{-Rang von } c_{ij}^{\cdot\cdot\ k}). \tag{116}$$

Diese Gruppe ist q-gliedrig und enthält alle infinitesimalen Transformationen, die sich als Klammerausdruck schreiben lassen, sie verschwindet also bei einer Abelschen Gruppe. Wird ein Vektor v^k in ξ dem Rande des Flächenelementes $f^{lj} d\sigma$ entlang (0)-parallel verschoben, so ist die Differenz zwischen Anfangs- und Endstellung bekanntlich

$$dv^k = f^{lj} R_{lji}^{\cdot\cdot\cdot\ k} v^i d\sigma = f^{lj} c_{lj}^{\ \ m} c_{mi}^{\cdot\cdot\ k} v^i d\sigma. \tag{117}$$

Die infinitesimalen Transformationen, die die infinitesimale Umgebung eines Punktes erfährt, wenn f alle möglichen Stellungen einnimmt, sind also gegeben durch die Gleichung

$$dv^k = e^m v^i c_{mi}^{\cdot\cdot\ k} dt, \tag{118}$$

wo e^m die ganze erste Ableitung der Gruppe durchläuft. Da die Gleichung (117) in jedem Punkte der X_r gilt und die $c_{ij}^{\ \ k}$ Konstanten sind, wird beim (0)-parallel Durchlaufen einer *endlichen* Schlinge die infinitesimale Umgebung in ξ eine *endliche* Transformation der ersten Ableitung erleiden. Die zu allen endlichen Schlingen gehörige Gruppe heißt nach Cartan die *Holonomiegruppe*[15]) der (0)-Übertragung, und es hat sich also herausgestellt, daß diese Holonomiegruppe mit der ersten Ableitung der adjungierten Gruppe identisch ist[15a]).

---

[15]) Les groupes d'holonomie des espaces généralisés, Acta Math. 28 (1925), p. 1–42.
[15a]) C. G. S. 77.

Aus c bilden wir jetzt die folgende Größenreihe:

$$\begin{aligned} g_i &= c_{ji}^{\cdot\cdot j}, \\ g_{ij} &= c_{ki}^{\cdot\cdot l} c_{lj}^{\cdot\cdot k}, \\ g_{ijk} &= g_{hi}^{\cdot\cdot l} g_{lj}^{\cdot\cdot m} g_{mk}^{\cdot\cdot h}, \\ &\cdots\cdots\cdots \end{aligned} \tag{119}$$

von der jedes Glied bei zyklischen Vertauschungen der Indizes invariant ist, z. B.

$$g_{ijk} = g_{jki} = g_{kij}. \tag{120}$$

Zu g_i gehört die $(r-1)$-gliedrige invariante Untergruppe

$$g_i e^i = 0. \tag{121}$$

In E_r nimmt sie die E_{r-1} in der $(r-1)$-Richtung von g_i ein. Da $g_i e^i$ eine Invariante ist, folgt aus (114)

$$c_{ij}^{\cdot\cdot k} g_k = 0, \tag{122}$$

eine Gleichung, die sich übrigens auch unmittelbar aus (119) und (13) ableiten läßt. Sie bringt zum Ausdruck, daß die erste Ableitung in der Gruppe von g_i enthalten ist. Auch folgt aus (122), daß das Feld des Vektors g_i wirbelfrei ist:

$$\overset{0}{\nabla}_{[\mu} g_{\lambda]} = \overset{+}{\nabla}_{[\mu} g_{\lambda]} + \frac{1}{2} c_{\mu\lambda}^{\cdot\cdot\nu} g_\nu = 0. \tag{123}$$

Außerdem folgt noch aus (122), daß

$$R_{\omega\mu\lambda}^{\cdot\cdot\cdot\lambda} = c_{\omega\mu}^{\cdot\cdot\alpha} c_{\alpha\lambda}^{\cdot\cdot\lambda} = 0, \tag{124}$$

und diese Gleichung bringt bekanntlich zum Ausdruck, daß die (0)-Übertragung inhaltstreu ist[16]). Die beiden anderen Übertragungen sind ebenfalls inhaltstreu, da sie integrabel sind[17]). Wir zeigen, daß die Beziehungen zwischen den drei Verschiebungen eines Inhalts gerade von dem Vektor g_λ abhängen. Betrachten wir dazu einen kontravarianten n-Vektor, der in der X_r über dx^ν das eine Mal $(+)$-äquipollent, das andere Mal (0)-äquipollent verschoben wird. Dann ist

$$0 = dv^{\nu_1 \cdots \nu_n} + \overset{+}{\Gamma}{}_{\lambda\mu}^{\lambda} v^{\nu_1 \cdots \nu_n} dx^\mu \tag{125}$$

bzw.

$$0 = dv^{\nu_1 \cdots \nu_n} + \overset{0}{\Gamma}{}_{\lambda\mu}^{\lambda} v^{\nu_1 \cdots \nu_n} dx^\mu, \tag{126}$$

---

[16]) R. K. S. 89.

[17]) Hurwitz, Göttinger Nachrichten 1897, S. 71, verwendet zur Invariantenbildung Volumenintegrale, deren Existenz auf der Inhaltstreue der $(-)$-Übertragung beruht.

und die Differenz der Endwerte beträgt also

$$(\overset{+}{\Gamma}{}^{\lambda}_{\lambda\mu} - \overset{0}{\Gamma}{}^{\lambda}_{\lambda\mu})\, v^{\nu_1 \dots \nu_n} dx^{\mu} = \frac{1}{2}\, g_{\mu}\, dx^{\mu}\, v^{\nu_1 \dots \nu_n}. \tag{127}$$

Nur im Falle, wo g_{μ} verschwindet, sind also die Verschiebungen eines Inhalts bei den drei Übertragungen gleich[18]).

Die erste merkwürdige Eigenschaft von g_{ij} ist, daß $c_{ij}^{\cdot\cdot l} g_{lk}$ ein *Trivektor*[19]) ist, wie sich leicht aus (119) und (13) ergibt. Die invariante Untergruppe dieses Trivektors

$$c_{ij}^{\cdot\cdot l}\, g_{lk}\, e^{k} = 0 \tag{128}$$

enthält offenbar die invariante Untergruppe von g_{ij}

$$g_{ij}\, e^{j} = 0\,. \tag{129}$$

Cartan hat bewiesen[20]), daß die invariante Untergruppe des Trivektors $c_{ij}^{\cdot\cdot l} g_{lk}$ alle integrablen invarianten Untergruppen enthält. Daraus folgt der Satz, daß eine Gruppe dann und nur dann integrabel ist, wenn der Trivektor $c_{ij}^{\cdot\cdot l} g_{lk}$ verschwindet.

Die zweite ist, daß die Gruppe dann und nur dann halbeinfach ist, wenn der Rang von g_{ij} gleich r ist[21]). Da dann auch der i-Rang und der k-Rang von $c_{ij}^{\cdot\cdot k}$ gleich r sein müssen, verschwindet g_i und die erste Ableitung fällt mit der Gruppe zusammen. Da auch die (0)-Ableitung von g_{ij} verschwindet, geht die Geometrie der (0)-Übertragung in eine Riemannsche Geometrie über. Aus (119) folgt

$$K_{\mu\lambda} = \frac{1}{4}\, g_{\mu\lambda}, \tag{130}$$

und es liegt also eine Einsteinsche Geometrie besonderer Art vor, in welcher die Ableitung der Krümmungsgröße verschwindet. Die E_r geht über in eine R_r, und die adjungierte Gruppe wird eine Gruppe von orthogonalen Transformationen. Da sich jetzt nicht nur c_{ijk} aus $c_{ij}^{\cdot\cdot k}$, sondern auch $c_{ij}^{\cdot\cdot k}$ aus dem Trivektor c_{ijk} berechnen läßt, ist die Klassifizierung der halbeinfachen Gruppen gleichbedeutend geworden mit der algebraischen Aufgabe, die Trivektoren in R_n zu klassifizieren, die der Bedingung

$$g^{ij}\, c_{i[kl}\, c_{m]jh} = 0 \tag{131}$$

genügen.

Die Größe g_{ijk} zerfällt infolge der zyklischen Invarianz in einen symmetrischen und einen alternierenden Teil:

$$g_{ijk} = g_{(ijk)} + g_{[ijk]}. \tag{132}$$

[18]) C. G. S. 69.

[19]) C & S I, S. 810.

[20]) Cartan, Thèse 1894, p. 109.

[21]) Cartan, Thèse 1894, p. 52.

Aus (13) folgt, daß der alternierende Teil der Hälfte des oben erwähnten Trivektors gleich ist

$$(133) \qquad 2g_{[ijk]} = c_{ij}^{\cdot\cdot l} g_{lk}.$$

Notwendige und hinreichende Bedingung für das Integrabelsein einer Gruppe ist also auch, daß $g_{[ijk]}$ verschwindet[22]). Für eine halbeinfache Gruppe verschwindet $g_{(ijk)}$ und wird $2g_{ijk}$ gleich c_{ijk}. Der alternierende Teil der Größe g_{ijkl} verschwindet infolge der zyklischen Invarianz. Für den Fall einer halbeinfachen Gruppe steht g_{ijkl} in Beziehung zur Krümmungsgröße, es ist

$$(134) \qquad g_{ijkl} - g_{ijlk} = K_{ijkl}$$

und

$$(135) \qquad g_{ijkl} = g_{(ijkl)} + \frac{1}{3} K_{ijkl} + \frac{1}{3} K_{lijk}.$$

Aus (135) folgt, daß $g_{ijkl} - g_{(ijkl)}$ die erste Größe der von Veblen und T. Y. Thomas[23]) aufgestellten Reihe der „normal tensors" ist. Diese Größenreihe entsteht bekanntlich, wenn man die kanonischen Parameter als Urvariablen wählt und dann die $\Gamma_{\lambda\mu}^{\nu}$ in einer Reihe nach ε entwickelt. Die höheren Glieder der Reihe (119) kamen bisher nicht zur Anwendung.

---

[22]) Cartan, Thèse 1894, p. 47.

[23]) The Geometry of paths, Trans. Amer. Math. Soc. **25** (1923), p. 551–608.

(Eingegangen am 28. 10. 1928.)

Über die Einheitengruppe eines endlichen Ringes[1]).

Von

Kenjiro Shoda in Berlin.

In einer früheren Arbeit[2]) habe ich den Begriff des *Strahls* eines im allgemeinen nichtkommutativen Ringes $\mathfrak{o}$ modulo einem Unterringe $\mathfrak{c}$ eingeführt. Unter dem Strahl $\{\mathfrak{o}, \mathfrak{c}\}$ verstehen wir nämlich die Gesamtheit der Elemente aus $\mathfrak{o}$, die modulo $\mathfrak{c}$ dem Einheitselement E von $\mathfrak{o}$ (dessen Existenz vorausgesetzt sei) kongruent sind. Die Gesamtheit der Einheiten (d. h. die Gesamtheit der Elemente mit reziproken Elementen) aus $\mathfrak{o}$ bildet durch Multiplikation eine (eigentliche) Gruppe, die ich die Einheitengruppe oder die zugehörige Gruppe von $\mathfrak{o}$ nenne. Ist $\mathfrak{c}$ ein Ideal[3]) in $\mathfrak{o}$, so bildet der Durchschnitt des Strahls $\{\mathfrak{o}, \mathfrak{c}\}$, der dann in der Form $\{\underline{\mathfrak{o} | \mathfrak{c}}\}$ geschrieben wird, und der zugehörigen Gruppe $\mathfrak{G}$ einen Normalteiler von $\mathfrak{G}$. Auf diese Weise entspricht einem Ideal ein Normalteiler, einem Restklassenringe eine Faktorgruppe und einer direkten Summe von Idealen ein direktes Produkt von Normalteilern.

In der vorliegenden Arbeit handelt es sich allgemeiner um Unter*ringe*, nicht Ideale. Im Fall der nilpotenten Unterringe kommen tatsächlich einfache Resultate: Der Strahl $\{\mathfrak{o}, \mathfrak{c}\}$ für einen nilpotenten Unterring $\mathfrak{c}$ bildet eine Untergruppe der zugehörigen Gruppe $\mathfrak{G}$ (Satz 2).

Als Vorbereitung für die Anwendung dieses Satzes beweise ich in § 1 nach einigen Bemerkungen über maximale nilpotente Unterringe, daß der *Durchschnitt aller maximalen nilpotenten Unterringe eines Ringes das maximale nilpotente Ideal* ist, wenn man die Existenz des Einheits-

[1]) Diese Arbeit wurde in Proc. of the Imperial Academy of Japan 5, Nr. 3, vorangezeigt.

[2]) K. Shoda, Über die Automorphismen einer endlichen Abelschen Gruppe, § 3, Math. Annalen **100** (1928), S. 674—686.

[3]) Unter einem Ideal soll im folgenden stets ein zweiseitiges Ideal verstanden werden.

elementes und das Bestehen des *Doppelkettensatzes*[4]) *für Rechtsideale* vorausgesetzt (Satz 1).

Ein *endlicher* Ring heißt ein p-Ring, wenn die Anzahl der Elemente eine Primzahlpotenz ist. Dann gilt der Satz, daß ein endlicher Ring stets die direkte Summe von p-Ringen ist (Satz 3). Ich betrachte daher einen p-Ring und beweise die folgenden Sätze: Der Strahl eines p-Ringes modulo einem maximalen nilpotenten Unterringe bildet eine Sylowgruppe[5]) der zugehörigen Gruppe (Satz 4). Der Strahl eines p-Ringes modulo dem maximalen nilpotenten Ideale bildet den Durchschnitt aller Sylowgruppen (Satz 6). Man kann dadurch leicht erkennen, daß einer Sylowgruppe eineindeutig ein maximaler nilpotenter Unterring entspricht. Man kann daher einige Sätze über Sylowgruppen in die Theorie der maximalen nilpotenten Unterringe eines endlichen Ringes übertragen und umgekehrt. Dadurch erhalte ich, daß *alle maximalen nilpotenten Unterringe eines endlichen Ringes miteinander konjugiert sind* (Satz 5).

Da der Automorphismenring[6]) einer endlichen Abelschen Gruppe die übliche Automorphismengruppe als seine zugehörige Gruppe enthält, so findet sich eine Anwendung unserer Sätze bei der Untersuchung der Sylowgruppen der Automorphismengruppe einer Abelschen Gruppe, deren Ordnung eine Primzahlpotenz ist, worauf ich aber jetzt nicht eingehe[7]).

§ 1.
Nilpotente Unterringe eines Ringes.

Es sei $\mathfrak{o}$ ein Ring mit Einheitselement E, $\mathfrak{n}$ ein nilpotentes Ideal in $\mathfrak{o}$ und $\mathfrak{m}'$ ein maximaler nilpotenter Unterring[8]) des Restklassenringes $\mathfrak{o}/\mathfrak{n}$, d. h. es gebe keinen nilpotenten Unterring von $\mathfrak{o}/\mathfrak{n}$, der $\mathfrak{m}'$ als einen echten Unterring enthält. Dabei soll unter einem nilpotenten Ring ein Ring $\mathfrak{w}$ verstanden werden, der der Bedingung $\mathfrak{w}^{\alpha} = 0$ für ein geeignetes α genügt. Die Gesamtheit der Elemente aus $\mathfrak{o}$, die in den Restklassen aus $\mathfrak{m}'$ enthalten sind, wird nach Herrn W. Krull[9]) mit $\mathfrak{n}^\frown\mathfrak{m}'$ bezeichnet.

---

[4]) E. Noether, Abstrakter Aufbau der Idealtheorie in algebraischen Zahl- und Funktionenkörpern, Math. Annalen **96** (1926), S. 26—61.

[5]) Unter einer Sylowgruppe der zugehörigen Gruppe eines p-Ringes verstehen wir nur eine solche, deren Ordnung eine Potenz von p ist, wenn die Anzahl der Elemente des p-Ringes eine Potenz von p ist.

[6]) A. Chatelet, Les groupes abéliens finis et les modules de points entiers, Lille 1925. K. Shoda, a. a. O.

[7]) H. A. Bender, Sylow subgroups in the groups of isomorphisms of prime power abelian groups, American Journal of Mathematics **45** (1923), S. 223—250. Dort werden die Sylowgruppen ohne Benützung des Automorphismenringes untersucht.

[8]) Die Existenz des maximalen nilpotenten Unterringes sei jetzt vorausgesetzt.

[9]) W. Krull, Theorie und Anwendung der verallgemeinerten Abelschen Gruppen, Sitzungsberichte der Heidelberger Akademie 1926.

Hilfssatz 1. $\mathfrak{m} = \mathfrak{n}^{-}\mathfrak{m}'$ *bildet einen maximalen nilpotenten Unterring von* $\mathfrak{o}$. *Umgekehrt umfaßt ein maximaler nilpotenter Unterring jedes nilpotente Ideal* $\mathfrak{n}$, *ist also von der Form* $\mathfrak{n}^{-}\mathfrak{m}'$, *wo* $\mathfrak{m}'$ *einen maximalen nilpotenten Unterring in* $\mathfrak{o}/\mathfrak{n}$ *bedeutet.*

Es ist klar, daß $\mathfrak{m}$ ein Ring ist. Ist nun $\mathfrak{n}^{\alpha} = 0$ und $\mathfrak{m}'^{\beta} = 0$, so ist $\mathfrak{m}^{\beta} = (\mathfrak{n}^{-}\mathfrak{m}')^{\beta}$ in $\mathfrak{n}$ enthalten, also ist $\mathfrak{m}^{\alpha\beta} = (\mathfrak{n}^{-}\mathfrak{m}')^{\alpha\beta} = 0$, d. h. $\mathfrak{m}$ ist ein nilpotenter Unterring. Ist $\mathfrak{m}$ in einem nilpotenten Unterringe $\mathfrak{v}$ enthalten, so ist $\mathfrak{m}'$ in $\mathfrak{v}/\mathfrak{n}$ enthalten, welches aber ein nilpotenter Unterring von $\mathfrak{o}/\mathfrak{n}$ ist. Nach der Definition von $\mathfrak{m}'$ ist also $\mathfrak{v} = \mathfrak{n}^{-}\mathfrak{m}'$, womit der erste Teil bewiesen ist. Umgekehrt umfaßt ein maximaler nilpotenter Unterring $\mathfrak{m}$ jedes nilpotente Ideal $\mathfrak{n}$. Denn die Summe $(\mathfrak{n}, \mathfrak{m})$ bildet einen Ring. Es ist ferner $(\mathfrak{n}, \mathfrak{m})^{r}$ für jedes r in $(\mathfrak{n}, \mathfrak{m}^{r})$ enthalten, da $\mathfrak{n}$ ein Ideal ist. Sind $\mathfrak{n}^{\alpha} = 0$, $\mathfrak{m}^{\beta} = 0$, so ist also $(\mathfrak{n}, \mathfrak{m})^{\alpha\beta} = 0$, d. h. die Summe bildet einen nilpotenten Unterring. Nach der Definition von $\mathfrak{m}$ ist also $\mathfrak{n}$ in $\mathfrak{m}$ enthalten, woraus ferner der zweite Teil des Satzes folgt.

Hilfssatz 2. *Es sei* $\mathfrak{o}$ *die direkte Summe* $\mathfrak{o}_1 + \mathfrak{o}_2 + \dots + \mathfrak{o}_t$ *von Idealen* $\mathfrak{o}_i$. *Ist* $\mathfrak{m}$ *ein maximaler nilpotenter Unterring von* $\mathfrak{o}$, *so ist* $\mathfrak{m}$ *die direkte Summe von solchen* $\mathfrak{m}_i$ *von* $\mathfrak{o}_i$. *Ist umgekehrt* $\mathfrak{m}_i$ *ein maximaler nilpotenter Unterring von* $\mathfrak{o}_i$, *so ist die direkte Summe* $\mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t$ *ein solcher von* $\mathfrak{o}$.

Es sei $E = E_1 + E_2 + \dots + E_t$ die entsprechende Zerlegung des Einheitselementes E, wo also E_i das Einheitselement von $\mathfrak{o}_i$ ist[10]). Es sei $\mathfrak{m}$ ein maximaler nilpotenter Unterring von $\mathfrak{o}$, der offenbar in $\mathfrak{m}E_1 + \mathfrak{m}E_2 + \dots + \mathfrak{m}E_t$ enthalten ist. Es ist aber $\mathfrak{m}E_1 + \mathfrak{m}E_2 + \dots + \mathfrak{m}E_t$ nilpotent, da $(\mathfrak{m}E_1 + \mathfrak{m}E_2 + \dots + \mathfrak{m}E_t)^{\alpha} = (\mathfrak{m}E_1)^{\alpha} + (\mathfrak{m}E_2)^{\alpha} + \dots + (\mathfrak{m}E_t)^{\alpha}$ $= \mathfrak{m}^{\alpha}E_1 + \mathfrak{m}^{\alpha}E_2 + \dots + \mathfrak{m}^{\alpha}E_t = 0$ ist, falls $\mathfrak{m}^{\alpha} = 0$ ist. Nach der Definition von $\mathfrak{m}$ ist also $\mathfrak{m} = \mathfrak{m}E_1 + \mathfrak{m}E_2 + \dots + \mathfrak{m}E_t$. Ist nun $\mathfrak{m}E_i$ in einem nilpotenten Unterringe $\mathfrak{m}_i$ von $\mathfrak{o}_i$ enthalten, so ist $\mathfrak{m}$ in $\mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t$ enthalten, der nilpotent ist, da $(\mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t)^{\varkappa} = \mathfrak{m}_1^{\varkappa} + \mathfrak{m}_2^{\varkappa} + \dots + \mathfrak{m}_t^{\varkappa}$ ist. Daher ist $\mathfrak{m}E_i$ nach der Definition von $\mathfrak{m}$ ein maximaler nilpotenter Unterring von $\mathfrak{o}_i$. Ist umgekehrt $\mathfrak{m}_i$ ein maximaler nilpotenter Unterring von $\mathfrak{o}_i$, so ist $\mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t$ ein solcher von $\mathfrak{o}$. Denn ist $\mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t$ in einem (maximalen) nilpotenten Unterring $\mathfrak{m}$ von $\mathfrak{o}$ enthalten, so ist $\mathfrak{m}_i$ in $\mathfrak{m}E_i$ enthalten, also ist nach der Definition von $\mathfrak{m}_i$ ferner $\mathfrak{m}_i = \mathfrak{m}E_i$ und $\mathfrak{m} = \mathfrak{m}_1 + \mathfrak{m}_2 + \dots + \mathfrak{m}_t$. Damit ist der Hilfssatz bewiesen.

Hilfssatz 3. *Ist* $\mathfrak{m}$ *ein nilpotenter Unterring von* $\mathfrak{o}$, *der der Bedingung* $\mathfrak{m}F = F\mathfrak{m}$ *für eine Einheit* F *aus* $\mathfrak{o}$ *genügt, so ist* $F + M$ *für jedes* M *aus* $\mathfrak{m}$ *eine Einheit.*

---

[10]) Fräulein E. Noether machte mich auf einen Fehler des ursprünglichen Beweises aufmerksam und gab mir diesen Beweis von Hilfssatz 2 an.

Es ist nämlich, falls $\mathfrak{m}^{2^e} = 0$ ist, $(F+M)(F-M_1)(F^2+M_2)\ldots(F^{2^{e-1}}+M_e) = F^{2^e} + M M_1 M_2 \ldots M_{e-1} = F^{2^e}$, wobei M_i durch $M M_1 M_2 \ldots M_{i-1} F^{2^{i-1}} = F^{2^{i-1}} M_i$ in $\mathfrak{m}^{2^i}$ bestimmt werden soll. Denn aus $\mathfrak{m} F = F \mathfrak{m}$ folgt auch $\mathfrak{m}^\lambda F^\mu = F^\mu \mathfrak{m}^\lambda$. Setzt man im Beweis $M_i = M^{2^{i-1}}$ für jedes i, so erhält man als einen speziellen Fall[11])

Hilfssatz 4. *Ist M ein nilpotentes Element, F eine mit M vertauschbare Einheit aus einem Ring, so ist $F+M$ eine Einheit.*

Wir setzen nun das Bestehen des *Doppelkettensatzes für Rechtsideale* in $\mathfrak{o}$ voraus. D. h. (*Teilerkettensatz*): Jede Kette von Rechtsidealen, bei der jedes Rechtsideal ein echter Oberring des vorangehenden ist, bricht im Endlichen ab. (*Vielfachenkettensatz*): Jede Kette von Rechtsidealen, bei der jedes Rechtsideal ein echter Unterring des vorangehenden ist, bricht im Endlichen ab.

Nach Herrn E. Artin[12]) kann man daraus die Existenz des maximalen nilpotenten Ideals schließen, welches alle nilpotenten Rechts- und Linksideale umfaßt. Der Restklassenring nach dem maximalen nilpotenten Ideal ist aber die direkte Summe von einfachen Ringen. Es sei $\mathfrak{n}$ das maximale nilpotente Ideal von $\mathfrak{o}$, $\mathfrak{o}/\mathfrak{n}$ die direkte Summe $\mathfrak{o}_1 + \mathfrak{o}_2 + \ldots + \mathfrak{o}_t$ von einfachen Ringen $\mathfrak{o}_i$. Ist $\mathfrak{m}_i$ ein maximaler nilpotenter Unterring von $\mathfrak{o}_i$, so ist $\mathfrak{m}' = \mathfrak{m}_1 + \mathfrak{m}_2 + \ldots + \mathfrak{m}_t$ nach Hilfssatz 2 ein solcher von $\mathfrak{o}/\mathfrak{n}$ und $\mathfrak{m} = \mathfrak{n}^\frown \mathfrak{m}'$ nach Hilfssatz 1 ein solcher von $\mathfrak{o}$ (und jeder maximale nilpotente Unterring von $\mathfrak{o}$ entsteht so).

Wir betrachten daher einen einfachen Ring, der nach einem Maclagan Wedderburnschen Satz[13]) mit einem vollständigen Matrizenringe $\mathfrak{a}$ vom Grade etwa n in einem (im allgemeinen nichtkommutativen) Körper K isomorph ist, d. h. es ist $\mathfrak{a} = \sum_{i,j} \mathsf{K} E_{ij}$, wo $E_{ij} E_{kl} = 0$ für $j \neq k$ und $= E_{il}$ für $j = k$, und E_{ij} mit jedem Elemente aus K vertauschbar ist. Ein maximaler nilpotenter Unterring $\mathfrak{b}$ von $\mathfrak{a}$ wird durch $\mathfrak{b} = \sum_{i<j} \mathsf{K} E_{ij}$ gegeben. Es ist klar, daß $\mathfrak{b}$ ein nilpotenter Unterring von $\mathfrak{a}$ ist. Ist P ein Element eines $\mathfrak{b}$ enthaltenden nilpotenten Unterringes $\mathfrak{c}$, so hat $\mathfrak{c}$ ein Element von der Gestalt $\sum_{i \geqq j} c_{ij} E_{ij}$, welches aber auch nilpotent sein muß. Also ist

---

11) Ist in den Hilfssätzen 3, 4 F ein Nullteiler, so ist $F+M$ ein Nullteiler, was gleichzeitig bewiesen ist. Ist F ein mit M vertauschbares nilpotentes Element, so ist $F+M$ wieder nilpotent. Denn aus $F^\alpha = 0$ und $M^\beta = 0$ folgt $(F+M)^{\alpha+\beta} = 0$.

12) E. Artin, Zur Theorie der hyperkomplexen Zahlen, § 1, Abhandlungen aus dem Math. Seminar zu Hamburg 5 (1927), S. 251—260.

13) Maclagan Wedderburn, On hypercomplex numbers, Proceedings of the London Math. Soc. 6 (1908). L. E. Dickson, Algebren und ihre Zahlentheorie, herausgegeben von A. Speiser, Zürich 1927, S. 120, 121. Das Bestehen dieses Satzes für den Ring mit Doppelkettensatz für Rechtsideale wurde von Herrn E. Artin a. a. O. Satz 11 bewiesen.

$c_{11} = c_{22} = \ldots = c_{nn} = 0$. Daher sind die Diagonalkoeffizienten[14]) von P gleich Null. Da P ein nilpotentes Element ist, so ist nach Hilfssatz 4 stets $cE + P$ eine Einheit, wo c ein von Null verschiedenes Element aus K bedeutet. Ist ein Koeffizient p_{i1} von P von Null verschieden, so setze man $c = p_{i1}$. Es gibt dann ein Element B aus $\mathfrak{b}$, so daß die i-te Zeile von $p_{i1}E + P + B$ und daher von $(p_{i1}E + P + B)\,X$ für jedes Element X aus $\mathfrak{a}$ mit der ersten Zeile derselben identisch ist. Also ist $(p_{i1}E + P + B)\,X$ für jedes X von E verschieden, d. h. $p_{i1}E + P + B$ ist keine Einheit. Dabei ist aber $P + B$ in $\mathfrak{c}$ enthalten, also gegen Hilfssatz 4 nilpotent. Also ist p_{i1} für jedes i gleich Null. Analog kann man behaupten, daß p_{ij} für jedes $i \geqq j$ gleich Null ist. Daher ist $\mathfrak{b} = \mathfrak{c}$ und $\mathfrak{b}$ ein maximaler nilpotenter Unterring von $\mathfrak{a}$. Einen maximalen nilpotenten $\mathfrak{m}$ eines allgemeinen Ringes $\mathfrak{o}$ kann man nun nach der obigen Bemerkung leicht konstruieren.

Ein anderer maximaler nilpotenter Unterring von $\mathfrak{a}$ wird durch $\mathfrak{b} = \sum\limits_{i>j} \mathsf{K} E_{ij}$ gegeben. Der Durchschnitt $[\mathfrak{b}, \mathfrak{b}]$ hat offenbar kein von Null verschiedenes Element. Nach Hilfssatz 2 erhält man

Hilfssatz 5. *Der Durchschnitt aller maximalen nilpotenten Unterringe eines halbeinfachen Ringes besitzt kein von Null verschiedenes Element.*

Nach diesem Hilfssatz erhält man weiter den folgenden allgemeinen Satz, der als ein Struktursatz des Ringes interessant ist:

Satz 1. *Der Durchschnitt aller maximalen nilpotenten Unterringe eines Ringes ist das maximale nilpotente Ideal des Ringes, wenn man die Existenz des Einheitselementes und das Bestehen des Doppelkettensatzes für Rechtsideale voraussetzt.*

Es sei nämlich $\mathfrak{n}$ das maximale nilpotente Ideal eines Ringes $\mathfrak{o}$. Man bilde nun den Restklassenring $\mathfrak{o}/\mathfrak{n}$, der bekanntlich halbeinfach ist. Der Durchschnitt aller maximalen nilpotenten Unterringe von $\mathfrak{o}/\mathfrak{n}$ ist nach Hilfssatz 5 ein Nullring. Daher folgt Satz 1 aus Hilfssatz 1 unmittelbar.

Es sei $\mathfrak{r}$ ein Unterring von $\mathfrak{o}$. Die Gesamtheit der Elemente P aus $\mathfrak{o}$ derart, daß $P\mathfrak{r}$ und $\mathfrak{r}P$ in $\mathfrak{r}$ enthalten ist, bildet einen Ring, den wir den *Normalisator* von $\mathfrak{r}$ nennen.

Ist $\mathfrak{p}$ der Normalisator eines maximalen nilpotenten Unterringes $\mathfrak{m}$, so besteht $\mathfrak{m}$ aus der Gesamtheit der nilpotenten Elemente aus $\mathfrak{p}$. Denn ist P bzw. M ein nilpotentes Element aus $\mathfrak{p}$ bzw. $\mathfrak{m}$, so ist das Produkt PM bzw. MP in $\mathfrak{m}$ enthalten. Der durch P und $\mathfrak{m}$ erzeugte Ring $\mathfrak{t}$ be-

[14]) Wir gebrauchen hier der Einfachheit halber die Ausdrucksweise der Matrizentheorie.

steht aus den Elementen $f(P)P+M$, wo $f(P)$ ein Polynom von P mit ganzen rationalen Koeffizienten bedeutet. Das Produkt von r Elementen aus $\mathfrak{t}$ hat dann die Gestalt $g(P)P^r+M$, also ist $\mathfrak{t}^a$ in $\mathfrak{m}$ enthalten, falls $P^a=0$, woraus folgt, daß $\mathfrak{t}$ nilpotent ist. Nach der Definition von $\mathfrak{m}$ ist also $\mathfrak{m}=\mathfrak{t}$ und P in $\mathfrak{m}$ enthalten, was zu beweisen war.

Daraus folgt nach dem Maclagan Wedderburnschen Satz

Hilfssatz 6. *Der Restklassenring $\mathfrak{p}/\mathfrak{m}$ ist die direkte Summe von Körpern, wenn der Doppelkettensatz für Rechtsideale in $\mathfrak{p}$ besteht.*

Ist ein Element P aus $\mathfrak{o}$ mit $\mathfrak{m}$ vertauschbar, d. h. $P\mathfrak{m}=\mathfrak{m}P$, und $\mathfrak{m}$ maximaler nilpotenter Unterring, *so ist P im Normalisator $\mathfrak{p}$ von $\mathfrak{m}$ enthalten.* Denn es ist $(P\mathfrak{m})^a=(\mathfrak{m}P)^a=\mathfrak{m}^aP^a=0$, falls $\mathfrak{m}^a=0$ ist, d. h. PM und MP für jedes M aus $\mathfrak{m}$ sind nilpotent. Jedes Element des durch $P\mathfrak{m}$ und $\mathfrak{m}$ erzeugten Ringes $\mathfrak{s}$ hat die Gestalt $P^\lambda M_0+P^{\lambda-1}M_1+\ldots+M_\lambda$, wo M_i ein Element aus $\mathfrak{m}$ bedeutet. Die Koeffizienten von P^i im Produkt von r Elementen aus $\mathfrak{s}$ sind wegen $P\mathfrak{m}=\mathfrak{m}P$ in $\mathfrak{m}^r$ enthalten, da aus $P\mathfrak{m}=\mathfrak{m}P$ auch $P^i\mathfrak{m}=\mathfrak{m}P^i$ folgt. Daher ist $\mathfrak{s}^a=0$. Nach der Definition von $\mathfrak{m}$ ist also $P\mathfrak{m}=\mathfrak{m}P$ in $\mathfrak{m}$ enthalten.

Um den Normalisator des oben konstruierten maximalen nilpotenten Unterringes eines Ringes zu bestimmen, beweisen wir

Hilfssatz 7. *Ist $\mathfrak{n}$ ein Ideal in $\mathfrak{o}$, $\mathfrak{m}'$ ein Unterring von $\mathfrak{o}/\mathfrak{n}$ und $\mathfrak{p}'$ der Normalisator von $\mathfrak{m}'$, so ist $\mathfrak{n}^\frown\mathfrak{p}'$ der Normalisator von $\mathfrak{n}^\frown\mathfrak{m}'$.*

Der Ring $\mathfrak{n}^\frown\mathfrak{m}'$ ist offenbar ein Ideal in $\mathfrak{n}^\frown\mathfrak{p}'$, da $\mathfrak{n}$ ein Ideal in $\mathfrak{o}$ ist. Ist $\mathfrak{n}^\frown\mathfrak{m}'$ ein Ideal in $\mathfrak{v}$, so ist $\mathfrak{m}'$ ein Ideal in $\mathfrak{v}/\mathfrak{n}$, also ist $\mathfrak{v}/\mathfrak{n}$ in $\mathfrak{p}'$ und daher $\mathfrak{v}$ in $\mathfrak{n}^\frown\mathfrak{p}'$ enthalten.

Klar ist nun

Hilfssatz 8. *Ist $\mathfrak{o}$ die direkte Summe $\mathfrak{o}_1+\mathfrak{o}_2+\ldots+\mathfrak{o}_t$ von Idealen $\mathfrak{o}_i$, $\mathfrak{m}_i$ ein Unterring von $\mathfrak{o}_i$ und $\mathfrak{p}_i$ der Normalisator von $\mathfrak{m}_i$ in $\mathfrak{o}_i$, so ist $\mathfrak{p}=\mathfrak{p}_1+\mathfrak{p}_2+\ldots+\mathfrak{p}_t$ der Normalisator von $\mathfrak{m}=\mathfrak{m}_1+\mathfrak{m}_2+\ldots+\mathfrak{m}_t$.*

Es sei wie früher $\mathfrak{n}$ das maximale nilpotente Ideal in $\mathfrak{o}$, $\mathfrak{o}/\mathfrak{n}$ die direkte Summe $\mathfrak{o}_1+\mathfrak{o}_2+\ldots+\mathfrak{o}_t$ von einfachen Ringen $\mathfrak{o}_i$, $\mathfrak{m}_i$ ein maximaler nilpotenter Unterring von $\mathfrak{o}_i$. Ist $\mathfrak{p}_i$ der Normalisator von $\mathfrak{m}_i$ in $\mathfrak{o}_i$, so ist $\mathfrak{p}'=\mathfrak{p}_1+\mathfrak{p}_2+\ldots+\mathfrak{p}_t$ der von $\mathfrak{m}'=\mathfrak{m}_1+\mathfrak{m}_2+\ldots+\mathfrak{m}_t$ und $\mathfrak{p}=\mathfrak{n}^\frown\mathfrak{p}'$ der von $\mathfrak{m}=\mathfrak{n}^\frown\mathfrak{m}'$. Daher betrachten wir wieder — wie es genügend ist — den einfachen Ring $\mathfrak{a}$. Der Normalisator des maximalen nilpotenten Unterringes $\mathfrak{b}=\sum_{i<j}\mathsf{K}E_{ij}$ von $\mathfrak{a}$ ist — wie man leicht sehen kann — gleich $\mathfrak{f}=\sum_{i\leqq j}\mathsf{K}E_{ij}$.

Hilfssatz 9. *Die zugehörige Gruppe $\mathfrak{F}$ von $\mathfrak{f}$ besteht aus den Elementen aus $\mathfrak{f}$ derart, daß der Koeffizient von E_{ii} für jedes i von Null verschieden ist.*

Denn ein Element F aus $\mathfrak{f}$ läßt sich durch $F = (\sum f_{ii} E_{ii})(E + \sum_{i<j} f_{ij} E_{ij})$ darstellen. Da der zweite Faktor nach Hilfssatz 4 stets eine Einheit ist, so ist F dann und nur dann eine Einheit, wenn der erste Faktor eine Einheit ist, woraus die Behauptung folgt.

Hilfssatz 10. *Die Gesamtheit der mit einem maximalen nilpotenten Unterringe $\mathfrak{m}$ vertauschbaren Einheiten bildet die zugehörige Gruppe (Einheitengruppe) des Normalisators $\mathfrak{p}$ von $\mathfrak{m}$, wenn man das Bestehen des Vielfachenkettensatzes für Rechtsideale in $\mathfrak{m}$ voraussetzt.*

Ist nämlich P eine in $\mathfrak{p}$ enthaltene Einheit von $\mathfrak{o}$, so ist $P\mathfrak{m}$ in $\mathfrak{m}$ enthalten. Ferner ist $P^i\mathfrak{m}$ in $P^{i-1}\mathfrak{m}$ enthalten. Es ist aber $P^i\mathfrak{m}$ ein Rechtsideal in $\mathfrak{m}$. Denn $P^i\mathfrak{m}\mathfrak{m}$ ist in $P^i\mathfrak{m}$ enthalten. Daher ist nach der Voraussetzung $P^\alpha\mathfrak{m} = P^{\alpha-1}\mathfrak{m}$ für ein geeignetes α, woraus aber $P\mathfrak{m} = \mathfrak{m}$ folgt, da $P^{\alpha-1}$ eine Einheit ist. Aus $P\mathfrak{m} = \mathfrak{m}$ folgt auch $\mathfrak{m} = P^{-1}\mathfrak{m}$ und daß $P^{-1}\mathfrak{m}P$ in $\mathfrak{m}$ enthalten ist. Also ist $\mathfrak{m}$ in $P\mathfrak{m}P^{-1}$ enthalten, und nach der Definition von $\mathfrak{m}$ ist $\mathfrak{m} = P\mathfrak{m}P^{-1}$ oder $P\mathfrak{m} = \mathfrak{m}P = \mathfrak{m}$, da $P\mathfrak{m}P^{-1}$ auch nilpotent ist. Damit ist Hilfssatz 10 bewiesen, da jedes mit $\mathfrak{m}$ vertauschbare Element nach oben in $\mathfrak{p}$ enthalten ist.

§ 2.

Sylowgruppen der Einheitengruppe eines endlichen Ringes.

Für einen Ring $\mathfrak{o}$ mit Einheitselement E gilt unabhängig vom Doppelkettensatz

Satz 2. *Der Strahl $\{\mathfrak{o}, \mathfrak{m}\}$ modulo einem nilpotenten Unterringe $\mathfrak{m}$ bildet eine Untergruppe der zugehörigen Gruppe (Einheitengruppe) $\mathfrak{G}$ von $\mathfrak{o}$.*[15])

Jedes Produkt zweier Elemente aus $\{\mathfrak{o}, \mathfrak{m}\}$ ist offenbar in $\{\mathfrak{o}, \mathfrak{m}\}$ enthalten. Nach dem Beweis von Hilfssatz 4, wobei jetzt F gleich dem Einheitselement E ist, bildet $\{\mathfrak{o}, \mathfrak{m}\}$ ferner eine (eigentliche) Gruppe, da dabei $(E - M)$, $(E + M^2)$, …, $(E + M^{2^\sigma-1})$ in $\{\mathfrak{o}, \mathfrak{m}\}$ enthalten sind; also ist jedes Element des Strahles $\{\mathfrak{o}, \mathfrak{m}\}$ eine Einheit, deren reziprokes Element auch in $\{\mathfrak{o}, \mathfrak{m}\}$ enthalten ist.

Wir betrachten nun einen endlichen Ring mit Einheitselement E. Ein endlicher Ring heiße ein p-Ring, wenn die Anzahl der Elemente des Ringes eine Primzahlpotenz ist.

Satz 3. *Ein endlicher Ring ist stets die direkte Summe von p-Ringen.*

Ein endlicher Ring $\mathfrak{o}$ läßt sich als eine additiv geschriebene Abelsche Gruppe bekanntlich in die direkte Summe $\mathfrak{o} = \mathfrak{o}_1 + \mathfrak{o}_2 + \dots + \mathfrak{o}_m$ zerlegen,

---

[15]) Ist der Ring endlich, so ist diese Untergruppe auflösbar. Zum Beweis vgl. K. Shoda, a. a. O., Beweis des Satzes 8.

wo die Anzahlen der Elemente aus $\mathfrak{o}_i$ teilerfremde Primzahlpotenzen $p_i^{e_i}$ sind. Dann besteht $\mathfrak{o}_i$ aus der Gesamtheit der Elemente P_i aus $\mathfrak{o}$, die der Bedingung $p_i^{e_i} P_i = 0$ genügen. Also bildet $\mathfrak{o}_i$ einen Ring. Wir haben also nur zu zeigen, daß $P_i P_j = 0$ ist, falls P_i bzw. P_j in $\mathfrak{o}_i$ bzw. $\mathfrak{o}_j$, $i \neq j$, enthalten ist. Es ist $p_i^{e_i} P_i P_j = p_j^{e_j} P_i P_j = 0$. Da p_i und p_j teilerfremd sind, so muß nach dem Distributivgesetz $P_i P_j = 0$ sein.

Wir betrachten nun einen einfachen p-Ring $\mathfrak{a}$. Der Ring $\mathfrak{a}$ ist dann mit einem vollständigen Matrizenringe vom Grade etwa n in einem Galoisschen Felde isomorph, da ein endlicher Körper stets nach einem Maclagan Wedderburnschen Satz[16]) kommutativ ist. Es sei p^m die Anzahl der Elemente in dem Galoisschen Felde, so ist die Anzahl der Elemente aus $\mathfrak{a}$ gleich p^{mn^2}.

Die Ordnung der zugehörigen Gruppe $\mathfrak{A}$ von $\mathfrak{a}$ ist dann gleich

$$(p^{mn}-1)(p^{mn}-p^m)\ldots(p^{mn}-p^{m(n-1)})$$
$$= p^{\frac{mn(n-1)}{2}}(p^{mn}-1)(p^{m(n-1)}-1)\ldots(p^m-1).^{17)}$$

Die Ordnung der Sylowgruppe von $\mathfrak{A}$ ist also gleich der Anzahl der Elemente des in § 1 konstruierten maximalen nilpotenten Unterringes $\mathfrak{b}$ in $\mathfrak{a}$.

Es sei nun $\mathfrak{n}$ das maximale nilpotente Ideal eines allgemeinen p-Ringes $\mathfrak{o}$, und zwar die Anzahl der Elemente aus $\mathfrak{n}$ gleich p^λ. Ist der Restklassenring $\mathfrak{o}/\mathfrak{n}$ die direkte Summe $\mathfrak{o}_1 + \mathfrak{o}_2 + \ldots + \mathfrak{o}_t$, wo $\mathfrak{o}_i$ mit dem vollständigen Matrizenringe des Grades n_i in einem Galoisschen Felde mit p^{m_i} Elementen isomorph ist, so ist die Anzahl der Elemente aus $\mathfrak{o}$ gleich $p^{\lambda + \sum_{i=1}^{t} m_i n_i^2}$ und die Ordnung der zugehörigen Gruppe $\mathfrak{G}$ gleich $p^{\lambda + \frac{1}{2}\sum_{i=1}^{t} m_i n_i (n_i - 1)} \prod_{i=1}^{t} (p^{m_i n_i}-1)(p^{m_i(n_i-1)}-1)\ldots(p^{m_i}-1)$, da bei der Strahlbildung dem Restklassenring $\mathfrak{o}/\mathfrak{n} = \mathfrak{o}_1 + \mathfrak{o}_2 + \ldots + \mathfrak{o}_t$ die Faktorgruppe $\mathfrak{G}/\mathfrak{N} = \mathfrak{G}_1 \times \mathfrak{G}_2 \times \ldots \times \mathfrak{G}_t$ entspricht, wo $\mathfrak{N}$ gleich dem Strahl $\{\mathfrak{o}\,|\,\mathfrak{n}\}$ und $\mathfrak{G}_i$ mit der zugehörigen Gruppe von $\mathfrak{o}_i$ isomorph ist[18]).

Satz 4. *Der Strahl* $\{\mathfrak{o}, \mathfrak{m}\}$ *eines p-Ringes* $\mathfrak{o}$ *modulo einem maximalen nilpotenten Unterringe* $\mathfrak{m}$ *ist eine Sylowgruppe der zugehörigen Gruppe (Einheitengruppe) von* $\mathfrak{o}$.

---

[16]) Maclagan Wedderburn, A theorem on finite algebra, Transactions of the American Math. Soc. 6; L. E. Dickson, On finite algebra, Göttinger Nachr. 1905; E. Artin, Über einen Satz von Herrn J. H. Maclagan Wedderburn, Abhandlungen aus dem Math. Seminar zu Hamburg 5 (1927), S. 245–250.

[17]) L. E. Dickson, Linear groups with an exposition of the Galois field theory, Leipzig 1901, S. 77.

[18]) Vgl. den Anfang der Einleitung und Hilfssatz 3.

Die Richtigkeit dieses Satzes für den in § 1 konstruierten maximalen nilpotenten Unterring $\mathfrak{m}$ eines p-Ringes $\mathfrak{o}$ erkennt man leicht, wenn man die Ordnung der Sylowgruppe und die Anzahl der Elemente des maximalen nilpotenten Unterringes vergleicht (siehe oben), da nach Satz 2 der Strahl modulo einem nilpotenten Unterringe nur aus Einheiten besteht und eine Gruppe bildet. Ist $\mathfrak{m}^*$ irgendein maximaler nilpotenter Unterring von $\mathfrak{o}$, so ist die Ordnung der Gruppe $\{\mathfrak{o}, \mathfrak{m}^*\}$ eine Potenz von p, und daher ist $\{\mathfrak{o}, \mathfrak{m}^*\}$ in einer Sylowgruppe enthalten, die sich aber durch eine Einheit P aus $\mathfrak{o}$ in der Form $P^{-1}\{\mathfrak{o}, \mathfrak{m}\}P = \{\mathfrak{o}, P^{-1}\mathfrak{m}P\}$ darstellt[19]). Also ist $\mathfrak{m}^*$ in $P^{-1}\mathfrak{m}P$ enthalten, und nach der Definition von $\mathfrak{m}^*$ ist daher $\mathfrak{m}^* = P^{-1}\mathfrak{m}P$, da $P^{-1}\mathfrak{m}P$ wieder ein nilpotenter Unterring ist. Also ist $\{\mathfrak{o}, \mathfrak{m}^*\} = P^{-1}\{\mathfrak{o}, \mathfrak{m}\}P$ eine Sylowgruppe der zugehörigen Gruppe $\mathfrak{G}$.

Gleichzeitig bewiesen ist der folgende Satz für einen p-Ring.

Satz 5. *Alle maximalen nilpotenten Unterringe eines endlichen Ringes sind miteinander konjugiert.*

Nach Satz 3 ist nämlich ein endlicher Ring $\mathfrak{o}$ die direkte Summe $\mathfrak{o}_1 + \mathfrak{o}_2 + \ldots + \mathfrak{o}_m$ von p-Ringen $\mathfrak{o}_i$. Ist $\mathfrak{m}$ ein maximaler nilpotenter Unterring von $\mathfrak{o}$, so ist nach Hilfssatz 2 also $\mathfrak{m} = \mathfrak{m}_1 + \mathfrak{m}_2 + \ldots + \mathfrak{m}_m$, wo $\mathfrak{m}_i$ ein maximaler nilpotenter Unterring von $\mathfrak{o}_i$ ist. Ist $\mathfrak{m}^* = \mathfrak{m}_1^* + \mathfrak{m}_2^* + \ldots + \mathfrak{m}_m^*$ ein anderer maximaler nilpotenter Unterring von $\mathfrak{o}$, so gibt es nach oben Einheiten P_i in $\mathfrak{o}_i$, so daß $P_i^{-1}\mathfrak{m}_i P_i = \mathfrak{m}_i^*$ ist. Setzt man $P = P_1 + P_2 + \ldots + P_m$, so ist $P^{-1} = P_1^{-1} + P_2^{-1} + \ldots + P_m^{-1}$ und $P^{-1}\mathfrak{m}P = \mathfrak{m}^*$, womit der Satz bewiesen ist.

Man kann also jeder Sylowgruppe der zugehörigen Gruppe $\mathfrak{G}$ eines p-Ringes $\mathfrak{o}$ eineindeutig durch Strahlbildung einen maximalen nilpotenten Unterring von $\mathfrak{o}$ zuordnen.

Aus Satz 4 folgt nach Satz 1

Satz 6. *Der Strahl* $\{\underline{\mathfrak{o} \,|\, \mathfrak{n}}\}$ *eines p-Ringes* $\mathfrak{o}$ *modulo dem maximalen nilpotenten Ideal* $\mathfrak{n}$ *ist der Durchschnitt aller Sylowgruppen der zugehörigen Gruppe (Einheitengruppe).*

Man kann nun einige Sätze über Sylowgruppen in die Theorie der maximalen nilpotenten Unterringe eines endlichen Ringes übertragen.

Zusatz 1. *Der Durchschnitt aller Sylowgruppen der zugehörigen Gruppe eines p-Ringes ist ein Normalteiler der zugehörigen Gruppe.*

Denn der Strahl $\{\underline{\mathfrak{o} \,|\, \mathfrak{n}}\}$ bildet einen Normalteiler von $\mathfrak{G}$, wenn $\mathfrak{n}$ ein nilpotentes Ideal in $\mathfrak{o}$ ist.

[19]) Vgl. etwa A. Speiser, Theorie der Gruppen von endlicher Ordnung, Berlin 1927, 2. Aufl., S. 66.

Zusatz 2. *Die im Normalisator eines maximalen nilpotenten Unterringes* $\mathfrak{m}$ *eines* p*-Ringes enthaltenen Einheiten bilden den Normalisator der Sylowgruppe* $\{\mathfrak{o}, \mathfrak{m}\}$ *der zugehörigen Gruppe.*

Dieser Satz folgt aus Hilfssatz 10 unmittelbar, da ein Element aus $\mathfrak{o}$ dann und nur dann mit $\{\mathfrak{o}, \mathfrak{m}\}$ vertauschbar ist, wenn es mit $\mathfrak{m}$ vertauschbar ist.

Zusatz 3. *Die Anzahl der maximalen nilpotenten Unterringe eines* p*-Ringes ist gleich der Anzahl der Sylowgruppen der zugehörigen Gruppe und wird durch* $\prod_{i=1}^{t}(p^{m_i n_i}-1)(p^{m_i(n_i-1)}-1)\ldots(p^{m_i}-1)/(p^{m_i}-1)^{n_i}$ *gegeben.*

Denn diese Anzahl ist gleich dem Index des Normalisators einer Sylowgruppe in der zugehörigen Gruppe. Die Ordnung des Normalisators $\mathfrak{P}$ einer Sylowgruppe $\mathfrak{M} = \{\mathfrak{o}, \mathfrak{m}\}$ ist gleich $p^{\lambda + \frac{1}{2}\sum_{i=1}^{t} m_i n_i(n_i-1)} \prod_{i=1}^{t}(p^{m_i}-1)^{n_i}$. Denn $\mathfrak{P}/\mathfrak{N} = \mathfrak{P}_1 \times \mathfrak{P}_2 \times \ldots \times \mathfrak{P}_t$, wo $\mathfrak{N}$ gleich dem Strahl $\{\underline{\mathfrak{o}|\mathfrak{n}}\}$ und $\mathfrak{P}_i$ der Normalisator der Sylowgruppe $\mathfrak{M}_i = \{\mathfrak{o}_i, \mathfrak{m}_i\}$ von $\mathfrak{G}_i$ ist. Die Ordnung von $\mathfrak{P}_i$ ist aber nach Hilfssatz 9 gleich $p^{\frac{1}{2} m_i n_i(n_i-1)}(p^{m_i}-1)^{n_i}$. Vergleicht man die Ordnung von $\mathfrak{P}$ mit der von $\mathfrak{G}$, so erkennt man die Richtigkeit von Zusatz 3.

Nach Satz 3 folgt aus Zusatz 3

Zusatz 4. *Die Anzahl der maximalen nilpotenten Unterringe eines endlichen Ringes ist kongruent* 1 *modulo* $(p_1, p_2, \ldots, p_s)$, *wobei die Anzahl der Elemente des Ringes aus* p_i *zusammengesetzt ist.*

D. h. die Anzahl stellt sich in der Form $1 + \sum_{i=1}^{s} r_i p_i$ dar.

(Eingegangen am 6. 2. 1929.)

Abriß einer arithmetischen Theorie der Galoisschen Körper.

(Zweite Mitteilung.)

Von

Öystein Ore in New Haven (Conn., U. S. A.).

In der ersten Mitteilung[1]) ist zunächst der einfachste, *reguläre* Fall der Zerlegungsgruppe untersucht worden, wo also höhere Verzweigungsgruppen nicht vorkommen. Weiter wurde in A, Kap. 2 die Grundlage für die Behandlung des *irregulären* Falles gegeben, indem gezeigt wurde, daß bei relativ-zyklischen Körpern vom Primzahlgrade p das volle Restsystem $(\bmod \mathsf{P}^\alpha)$, wo P ein Primidealteiler von p ist, immer durch Adjunktion einer Wurzel einer binomischen oder einer trinomischen Normalkongruenz $(\bmod \mathsf{P}^\alpha)$ erhalten werden kann. Die Form der Normalkongruenzen hängt nur von der Relativdifferente des Körpers ab.

In dieser zweiten Mitteilung werden nun diese Resultate für den Aufbau einer Theorie des irregulären Falles angewandt, welche wieder zum Studium der Zerlegungsgruppe dient.

Anstatt der Reihe der Verzweigungsgruppen wird zuerst eine Kompositionsreihe der Zerlegungsgruppe betrachtet; dementsprechend erhält man eine Reihe von relativ-zyklischen Körpern vom Relativgrade p, welche ich Irregularkörper genannt habe und welche die Reihe der Verzweigungskörper enthält. Daraus folgt der Hauptsatz, daß man den vollständigen Restbereich $(\bmod \mathsf{P}^\alpha)$ durch sukzessive Adjunktionen der Wurzeln von binomischen und trinomischen Normalkongruenzen von der Form

$$x^p - \frac{1}{r_i}\tau^{b_i}\pi_{i-1}^{e_i+1}x^{r_i} - \beta_{i-1}\pi_{i-1} \equiv 0 \pmod{\mathsf{P}^\alpha}$$

oder

$$x^p - \beta_{i-1}\pi_{i-1} \equiv 0 \pmod{\mathsf{P}^\alpha}$$

[1]) Math. Annalen 100 (1928), S. 650—673. Diese Abhandlung wird im folgenden kurz mit A bezeichnet.

aufbauen kann. Weiter werden die Verzweigungen μ_i [2]) durch die Konstanten der Normalkongruenzen bestimmt, woraus sofort unter Anwendung der Dedekind-Henselschen Ungleichung obere Grenzen für die μ_i angegeben werden können. Als Schlußstein von Kap. 1 wird bewiesen, daß die Normalkongruenzen für ein gegebenes Primideal entweder alle binomisch oder alle trinomisch sind, und daß im trinomischen Falle der Exponent $r_i = r$ eine von i unabhängige Konstante ist. Als eine Anwendung hiervon folgt für die Verzweigungen

$$\mu_1 \equiv \mu_2 \equiv \ldots \equiv 1 - r \pmod{p},$$

worin der Satz von Speiser [3]) enthalten ist, daß alle Verzweigungen einander $(\operatorname{mod} p)$ kongruent sind.

In Kap. 2 wird diese allgemeine Theorie zu einer vollständigen Untersuchung des binomischen Falles angewandt. In diesem Falle ist

$$e_0 \equiv 0 \pmod{p-1}, \qquad \mu_i = \frac{e_0 p^i}{p-1} + 1 \qquad (i = 1, 2, \ldots, s)$$

und die Verzweigungsgruppe ist zyklisch. Die vollständige Struktur der Trägheitsgruppe (Satz 16) und Zerlegungsgruppe (Satz 17) wird bestimmt. Die Form der Zerlegungsgruppe ist verhältnismäßig kompliziert, aber gruppentheoretisch interessant. Im binomischen Falle wird die Trägheitsgruppe nur für $p = 2$ Abelsch.

In einer letzten Mitteilung sollen verschiedene andere Fragen der arithmetischen Theorie der Galoisschen Körper behandelt werden, speziell wird der trinomische Fall eingehender studiert.

Kapitel 1.
Normalkongruenzen für die Irregularkörper.

§ 1.
Einführung der Irregularkörper.

Für den regulären Fall ist in A, Kap. 1 die vollständige Beziehung zwischen Gruppeneigenschaften und Gleichungseigenschaften aufgestellt. Dieselbe Aufgabe soll nun in dem schwierigeren Falle behandelt werden, wo auch höhere Verzweigungsgruppen vorkommen.

Den Galoisschen Körper K erhält man (vgl. A, Kap. 1, § 1) aus dem Regularkörper K_0 (= erster Verzweigungskörper) dadurch, daß man durch sukzessive Adjunktionen die Reihe der höheren Verzweigungskörper

(1) $$K_0 = K_{V_1}, K_{V_2}, \ldots, K_{V_{k+1}} = K$$

[2]) Man vgl. A, Kap. 1, § 1.

[3]) A. Speiser, Die Zerlegungsgruppe, Journ. f. Math. **149** (1919), S. 174–188. Diese Arbeit wird im folgenden als Speiser zitiert.

aufbaut. Hier ist allgemein K_{V_i} ein Relativkörper vom Grade $p^{s_{i-1}}$ zu $K_{V_{i-1}}$, wobei also

(2) $$s_1 + s_2 + \ldots + s_k = s,$$

wenn die Ordnung e des betrachteten Primideals P die Form $e = e_0 p^s$, $(e_0, p) = 1$ hat. Weiter sei

(3) $$G_0 = G_{V_1}, G_{V_2}, \ldots, G_{V_{k+1}} = 1$$

die Reihe der entsprechenden Verzweigungsgruppen, wo bekanntlich die Faktorgruppe $G_{V_{i-1}}/G_{V_i}$ Abelsch und vom Typus $\overbrace{p, p, \ldots, p}^{s_{i-1}}$ ist. Wie früher werden die Zerlegungsgruppe und Trägheitsgruppe mit G_Z und G_T bezeichnet.

Im folgenden erweist es sich nun sehr oft als vorteilhaft, nicht mit den Gruppen (1), sondern mit einer Kompositionsreihe der Gruppe G_0

(4) $$(G_{V_1} =) G_0, G_1, G_2, \ldots, G_s = 1$$

zu operieren, wobei allgemein die Gruppe G_i die i-te *Irregulargruppe* heißen soll. Die Faktorgruppe G_{i-1}/G_i ist zyklisch von der Ordnung p und die Gruppe G_i hat folglich die Ordnung p^{s-i}. Entsprechend (4) erhält man eine Reihe von s *Irregularkörpern*

(5) $$(K_{V_1} =) K_0, K_1, \ldots, K_s = K,$$

wobei K_i ein zyklischer Relativkörper vom Relativgrade p zu K_{i-1} ist. Der Relativgrad zu K_T von K_i wird $e_0 p^i$, während der Galoissche Körper K ein Relativkörper vom Grade p^{s-i} zu K_i wird.

Die Kompositionsreihe (4) wird in der folgenden Weise definiert: Es sei S_1 eine Substitution in G_{V_1}, welche nicht zu G_{V_2} gehört. Wenn dann π eine Primzahl in bezug auf P in K ist, so folgt nach der Definition von G_{V_1}

(6) $$S_1 : \pi \equiv \pi + \omega_1 \pi^{\mu_1} \pmod{\mathrm{P}^{\mu_1+1}},$$

wo μ_1 die in A, Kap. 1, § 1 definierte Verzweigung von G_{V_1} ist. Durch Wiederholung erhält man aus (6)

$$S_1^r : \pi \equiv \pi + r\omega_1 \pi^{\mu_1} \pmod{\mathrm{P}^{\mu_1+1}},$$

so daß speziell S_1^p eine Substitution in G_{V_2} ist.

Es folgt nun leicht, daß man genau s_1 (mod P) linear unabhängige Zahlen

(7) $$\omega_1, \omega_2, \ldots, \omega_{s_1}$$

so bestimmen kann, daß, wenn S eine beliebige Substitution in G_{V_1} ist, wird

$$S : \pi \equiv \pi + a\pi^{\mu_1} \pmod{\mathrm{P}^{\mu_1+1}},$$

wobei α die Form

$$\alpha = r_1 \omega_1 + r_2 \omega_2 + \dots + r_{s_1} \omega_{s_1}$$

mit ganzen rationalen r_i hat. Wenn daher S_i eine Substitution ist, für die

$$S_i : \pi \equiv \pi + \omega_i \pi^{\mu_1} \pmod{\mathrm{P}^{\mu_1+1}},$$

so kann man die ganze Verzweigungsgruppe G_{V_1} in der Form

$$G_{V_1} = S_1^{r_1} S_2^{r_2} \dots S_{s_1}^{r_{s_1}} G_{V_2} \qquad (r_i = 0, 1, \dots, p-1)$$

darstellen, wobei immer S_i^p eine Substitution in G_{V_2} ist, und weiter

$$S_i S_j = S_j S_i S^{(2)},$$

wo $S^{(2)}$ in G_{V_2} liegt.

Diejenigen Substitutionen in G_{V_1}, welche S_1 nicht enthalten, bilden eine Gruppe p^{s-1}-ter Ordnung, und diese ist die erste Irregulargruppe G_1. Die Gruppe G_1 ist offenbar ein Normalteiler von G_0 und G_0/G_1 ist zyklisch von der Ordnung p. Ebenso bilden diejenigen Substitutionen, welche weder S_1 noch S_2 enthalten, die zweite Irregulargruppe G_2 von der Ordnung p^{s-2} usw. Durch eine ähnliche Behandlung der nächsten Verzweigungsgruppen erhält man leicht:

Satz 1. ***Man kann in der ersten Verzweigungsgruppe s Substitutionen***

$$S_1, S_2, \dots, S_s \tag{8}$$

so bestimmen, daß die Gruppe in der Form

$$G_0 = S_1^{r_1} S_2^{r_2} \dots S_s^{r_s} \qquad (r_j = 0, 1, \dots, p-1)$$

dargestellt werden kann. Die i-te Irregulargruppe ist dann durch

$$G_i = S_{i+1}^{r_i} \dots S_s^{r_s} \qquad (r_j = 0, 1, \dots, p-1) \tag{9}$$

definiert, wobei für alle i immer S_{i+1}^p in einer Gruppe G_j $(j \geqq i+1)$ enthalten ist.

Es ist einleuchtend, daß in (8) die s_1 ersten Substitutionen S_i in G_{V_1}, die s_2 nächsten in G_{V_2} usw. liegen. Für die Irregularkörper K_i, welche den Irregulargruppen (9) entsprechen, werden wir der Bequemlichkeit wegen die kürzere Ausdrucksweise anwenden, daß K_i *zwischen den Verzweigungskörpern K_{V_r} und $K_{V_{r+1}}$ liegt*, wenn K_i ein (echter oder unechter) Unterkörper von $K_{V_{r+1}}$, aber nicht von K_{V_r} ist.

Eine Verzweigungsgruppe ist bekanntlich ein Normalteiler von allen vorangehenden Verzweigungsgruppen, sowie von der Trägheitsgruppe G_T und Zerlegungsgruppe G_Z. Eine Irregulargruppe ist, wie man leicht sieht, ein Normalteiler von allen vorangehenden Irregulargruppen, aber allgemein nicht von G_T und G_Z. Wenn S_{i+1} eine Substitution ist, für die

$$S_{i+1} : \pi \equiv \pi + \omega \pi^{\mu_r} \pmod{\mathrm{P}^{\mu_r+1}}$$

ist, so folgt leicht mit den Bezeichnungen in A, Kap. 1, § 5 [4])

$$(10) \qquad T^{-1} S_{i+1} T : \pi \equiv \pi + \omega \tau^{b_0 p^{-s}(\mu_r - 1)} \pi^{\mu_r} \pmod{\mathsf{P}^{\mu_r+1}}$$

$$(11) \qquad Z^{-1} S_{i+1} Z : \pi \equiv \pi + \omega^p \tau^{\lambda_0 p^{-s}(\mu_r - 1)} \pi^{\mu_r} \pmod{\mathsf{P}^{\mu_r+1}},$$

wo

$$\lambda_0 = \frac{a_0(p-1)}{e_0}, \qquad b_0 = \frac{\mathsf{p}^f - 1}{e_0}.$$

§ 2.

Normalkongruenzen für die Irregularkörper.

Im folgenden sollen nun die Eigenschaften der Irregularkörper eingehender studiert werden. Nach einer Bemerkung in A, Kap. 1, § 3 folgt sofort, daß wenn P_0 das Primideal im Regularkörper bezeichnet, worin P aufgeht, so besteht in den Irregularkörpern eine Zerlegung

$$\mathsf{P}_0 = \mathsf{P}_1^p = \mathsf{P}_2^{p^2} = \ldots = \mathsf{P}_s^{p^s} = \mathsf{P}^{p^s},$$

wobei jedes Primideal P_i den Relativgrad 1 hat. Da nun K_i nach der Definition ein relativ-zyklischer Körper vom Relativgrade p zu K_{i-1} ist, so kann man die Resultate aus A, Kap. 2, § 4 anwenden. Man erhält daraus eine Reihe von Tatsachen, welche in dem folgenden Hauptsatz zusammengefaßt werden können:

Satz 2. *Wenn π_{i-1} eine Primzahl in K_{i-1} in bezug auf P_{i-1} ist, so kann man immer eine Primzahl π_i in K_i in bezug auf P_i so bestimmen, daß π_i für ein beliebig hohes α entweder einer binomischen Kongruenz*

$$(12) \qquad x^p - \pi_{i-1}\beta_{i-1} \equiv 0 \pmod{\mathsf{P}^\alpha} \qquad (i = 1, 2, \ldots, s)$$

oder einer trinomischen Kongruenz

$$(13) \qquad x^p - \frac{1}{r_i}\tau^{b_i}\pi_{i-1}^{c_i+1} x^{r_i} - \pi_{i-1}\beta_{i-1} \equiv 0 \pmod{\mathsf{P}^\alpha} \qquad (i = 1, 2, \ldots, s)$$

genügt. Der binomische oder trinomische Fall tritt ein, je nachdem die relative Supplementzahl ϱ_i von P_i in bezug auf K_{i-1} durch

$$(14) \qquad \varrho_i = e_0 p^i$$

oder

$$(15) \qquad \varrho_i = c_i p + r_i < e_0 p^i \qquad (1 \leqq r_i \leqq p-1)$$

gegeben ist; dabei ist also bekanntlich ϱ_i dadurch definiert, daß die Relativdifferente von K_i in bezug auf K_{i-1} genau durch $\mathsf{P}_i^{p-1+\varrho_i}$ teilbar sein

[4]) Entsprechende Formeln kommen bei Speiser vor. Vgl. Speiser, Formel I u. II, § 3.

soll. In jedem Falle ist

$$(16)\qquad \varrho_i \equiv 0 \pmod{p-1}$$

und weiter im trinomischen Falle

$$(17)\qquad b_i \equiv 0 \pmod{p-1}.$$

Die Zahl β_{i-1}, welche also in K_{i-1} liegt, kann immer in der Form

$$(18)\qquad \beta_{i-1} = 1 + \tau^{a_1}\pi_{i-1} + \ldots + \tau^{a_{\frac{\varrho_i}{p-1}}}\pi_{i-1}^{\frac{\varrho_i}{p-1}}$$

geschrieben werden.

§ 3.
Bestimmung der Verzweigungen.

Die Normalkongruenzen des Satzes 2 sind nun die wichtigsten Hilfsmittel für das Studium der arithmetischen Eigenschaften des Galoisschen Körpers.

Die erste Aufgabe, welche hier behandelt werden soll, ist die Bestimmung des Zusammenhanges zwischen den Supplementzahlen ϱ_i und den in A, Kap. 1, § 1 definierten Verzweigungen μ_r. Die Verzweigung μ_r einer Verzweigungsgruppe G_{V_r} war dadurch festgelegt, daß wenn S eine Substitution in G_{V_r}, π eine Primzahl in bezug auf P bedeutet, so ist

$$S:\pi \equiv \pi + \omega\pi^{\mu_r} \pmod{\mathsf{P}^{\mu_r+1}},$$

wo ω nicht durch P teilbar ist.

Aus A, Kap. 1, § 5 folgt, daß eine beliebige Primzahl π in K einer irreduziblen Kongruenz

$$(19)\qquad F_0(x) = x^{p^s} + \pi_0 C_1^{(0)} x^{p^s-1} + \ldots + \pi_0 C_{p^s}^{(0)} \equiv 0 \pmod{\mathsf{P}^a}$$

im Regularkörper genügt, und die Zahl $F_0'(\pi)$ wird genau dieselbe Potenz von P enthalten wie die Relativdifferente von K in bezug auf K_0.

Es seien nun

$$(20)\qquad \pi = \pi_1,\ \pi_2, \ldots, \pi_{p^s}$$

die verschiedenen Primzahlen, welche man erhält, wenn man auf π die p^s Substitutionen der Regulargruppe G_0 anwendet. Diese sind natürlich alle Wurzeln von (19) und man erhält daraus sofort

$$(21)\qquad F_0'(\pi) \equiv (\pi-\pi_2)(\pi-\pi_3)\ldots(\pi-\pi_{p^s}) \pmod{\mathsf{P}^a}.$$

In (21) sind aber $p^s - p^{s-s_1}$ Faktoren genau durch P^{μ_1}, weiter $p^{s-s_1} - p^{s-s_1-s_2}$ genau durch P^{μ_2} teilbar usw. und die Relativdifferente wird folglich genau durch P^{Δ_0} teilbar, wo

$$(22)\qquad \Delta_0 = (p^s - p^{s-s_1})\mu_1 + (p^{s-s_1} - p^{s-s_1-s_2})\mu_2 + \ldots + (p^{s_k}-1)\mu_k.$$

Da die Differente des Regularkörpers genau durch $\mathsf{P}^{p^s(e_0-1)}$ teilbar ist, erhält man den bekannten Hilbertschen Satz;

Satz 3. *Die Differente des Galoisschen Körpers K ist genau durch P^{Δ} teilbar, wo*

$$\Delta = e - p^s + (p^s - p^{s-s_1})\mu_1 + \ldots + (p^{s_k} - 1)\mu_k.$$

Unsere Aufgabe war aber speziell die Relativdifferente eines Irregularkörpers K_i in bezug auf K_{i-1} zu berechnen. Betrachtet man zunächst den ersten Irregularkörper K_1, so folgt leicht nach der eben angewandten Methode, daß die Relativdifferente von K in bezug auf K_1 genau durch P^{Δ_1} teilbar ist, wo

$$\Delta_1 = (p^{s-1} - p^{s-s_1})\mu_1 + \ldots + (p^{s_k} - 1)\mu_k,$$

und allgemein beweist man, daß die Relativdifferente von K in bezug auf K_i genau durch P^{Δ_i} teilbar ist, wo

$$\Delta_i = (p^{s-i} - p^{s-s_1-\ldots-s_v})\mu_v + \ldots + (p^{s_k} - 1)\mu_k,$$

wenn K_i zwischen K_{V_v} und $K_{V_{v+1}}$ liegt.

Daraus folgt aber sofort nach einem bekannten Satz über Differenten, daß die Relativdifferente von K_i in bezug auf K_{i-1} genau durch

$$\mathsf{P}^{\Delta_{i-1}-\Delta_i} = \mathsf{P}^{p^{s-i}(p-1)\mu_v}$$

teilbar ist, und man hat den Satz:

Satz 4. *Wenn der Irregularkörper K_i zwischen K_{V_v}* und *$K_{V_{v+1}}$ liegt, so ist die Relativdifferente von K_i in bezug auf K_{i-1} genau durch $\mathsf{P}_i^{(p-1)\mu_v}$ teilbar.*

Dieser Satz gibt sofort die Relation zwischen den Verzweigungen μ_v und den Supplementzahlen ϱ_i. Die Relativdifferente von K_i zu K_{i-1} ist nämlich andererseits genau durch $P_i^{p-1+\varrho_i}$ teilbar, also nach Satz 4

$$p - 1 + \varrho_i = (p-1)\mu_v$$

und nach (16)

$$\mu_v = \frac{\varrho_i}{p-1} + 1.$$

Satz 5. *Wenn K_i zwischen K_{V_v} und $K_{V_{v+1}}$ liegt, besteht zwischen der Verzweigung μ_v und der relativen Supplementzahl ϱ_i die Beziehung*

(23) $$\mu_v = \frac{\varrho_i}{p-1} + 1.$$

Für alle Körper K_i zwischen K_{V_v} und $K_{V_{v+1}}$ hat daher ϱ_i denselben Wert, und die Reihe der Zahlen

(24) $$\frac{\varrho_i}{p-1} + 1 \qquad (i = 1, 2, \ldots, s)$$

stimmt mit der Reihe

$$(25)\qquad \overbrace{\mu_1, \dots \mu_1}^{s_1}, \quad \overbrace{\mu_2, \dots \mu_2}^{s_2}, \dots \overbrace{\mu_k, \dots \mu_k}^{s_k}$$

überein.

Aus Satz 5 fließen schon verschiedene wichtige Eigenschaften des Körpers. Die Gleichung (7) in A, Kap. 1 zeigt sofort, daß

$$(26)\qquad \varrho_1 \leqq \varrho_2 \leqq \varrho_3 \leqq \cdots \leqq \varrho_s.$$

Weiter leitet man aber unter Anwendung der Dedekind-Henselschen Ungleichung obere Grenzen für die Verzweigungen ab, welche das Resultat von Speiser[8]) verschärfen.

Aus (14) und (15) folgt nämlich $\varrho_i \leqq e_0 p^i$, und aus der Identität der beiden Reihen (24) und (25) ergibt sich dann

$$\mu_1 \leqq \left[\frac{e_0 p}{p-1}\right] + 1, \qquad \mu_2 \leqq \left[\frac{e_0 p^{s_1+1}}{p-1}\right] + 1, \dots$$

und es besteht der allgemeine Satz:

Satz 6. *Für die Verzweigungen μ_r hat man die obere Begrenzung*

$$(27)\qquad \mu_r \leqq \left[\frac{e_0 p^{s_1+\dots+s_{r-1}+1}}{p-1}\right] + 1 \qquad (r = 1, 2, \dots, k).$$

Speziell ist also für die größte Verzweigung

$$\mu_k \leqq \left[\frac{e_0 p^{s-s_k+1}}{p-1}\right] + 1 = \left[\frac{e}{(p-1)p^{s_k-1}}\right] + 1,$$

also sicher

$$(28)\qquad \mu_k \leqq \left[\frac{e}{p-1}\right] + 1,$$

wie von Herrn Speiser bewiesen.

§ 4.

Einteilung der Normalkongruenzen.

Es soll nun gezeigt werden, daß das System der Normalkongruenzen, welche nach Satz 2 das vollständige Restsystem (mod P^a) in K definieren, eine ganz spezielle und einfache Form haben muß. Dies folgt aus dem folgenden wichtigen Satz, der in diesem Paragraphen bewiesen werden soll:

Satz 7. *In der Reihe der Normalkongruenzen, welche nach Satz 2 die sukzessiven Irregularkörper definieren, sind entweder alle Kongruenzen binomisch*

$$(29)\qquad x^p - \pi_{i-1}\beta_{i-1} \equiv 0 \pmod{\mathsf{P}^a} \qquad (i = 1, 2, \dots, s)$$

[8]) Speiser, § 3, S. 183.

oder alle trinomisch

(30) $$x^p - \frac{1}{r}\tau^{b_i}\pi_{i-1}^{e_i+1}x^r - \pi_{i-1}\beta_{i-1} \equiv 0 \pmod{\mathsf{P}^a} \qquad (i = 1, 2, \ldots, s),$$

wobei der Exponent r von i unabhängig ist.

Durch diesen Satz kann man also immer die Untersuchungen in zwei Fälle zerlegen, je nachdem das Primideal P einer Kette von binomischen oder einer Kette von trinomischen Kongruenzen entspricht.

Um den Satz 7 zu beweisen, wird angenommen, daß π_i eine Primzahl ist, welche der Kongruenz (12) oder (13) genügt. Nach A, Kap. 2, § 4 haben dann die übrigen Lösungen die Form

(31) $$\pi_i + \tau^a\pi_i^{\mu_\nu}, \quad \pi_i + 2\,\tau^a\pi_i^{\mu_\nu}, \ldots, \pi_i + (p-1)\,\tau^a\pi_i^{\mu_\nu} \pmod{\mathsf{P}_i^{\mu_\nu+1}},$$

wo also wie früher vorausgesetzt wird, daß K_i zwischen K_{V_ν} und $K_{V_{\nu+1}}$ liegt. Man erhält die Wurzeln (31) aus π_i, indem man wiederholt die Substitution S_i in G_{i-1} anwendet.

Aus π_i wird nun die Primzahl π_{i+1} in K_{i+1} abgeleitet, wo also π_{i+1} entweder einer Kongruenz

(32) $$x^p - \beta_i\pi_i \equiv 0 \pmod{\mathsf{P}^a}$$

oder einer Kongruenz

(33) $$x^p - \frac{1}{r_{i+1}}\tau^{b_{i+1}}\pi_i^{e_{i+1}+1}x^{r_{i+1}} - \pi_i\beta_i \equiv 0 \pmod{\mathsf{P}^a}$$

genügt.

Der Körper K_{i+1} liegt entweder zwischen K_{V_ν} und $K_{V_{\nu+1}}$ so wie K_i, oder zwischen $K_{V_{\nu+1}}$ und $K_{V_{\nu+2}}$. Im ersten Falle ist aber nach Satz 5 $\varrho_i = \varrho_{i+1}$, so daß die Kongruenzen für π_i und π_{i+1} gleichzeitig binomisch oder trinomisch sein müssen; im trinomischen Falle ist dann auch offenbar $r_i = r_{i+1}$, wie bewiesen werden sollte.

Um den Satz 7 zu beweisen, ist es daher nur notwendig den Fall zu betrachten, wo K_{i+1} zwischen $K_{V_{\nu+1}}$ und $K_{V_{\nu+2}}$ liegt.

Zuerst wird vorausgesetzt, daß π_i einer binomischen Kongruenz genügt, und es soll gezeigt werden, daß dann auch π_{i+1} einer binomischen Kongruenz genügen muß.

Nimmt man nämlich an, π_{i+1} genüge der trinomischen Kongruenz (33), so wendet man auf π_{i+1} die Substitution S_i in G_{i-1} an und erhält eine neue Primzahl $\pi'_{i+1} = S_{i-1} : \pi_{i+1}$. Da nun die Gruppe G_{i+1} ein Normalteiler von G_{i-1} ist, gehört auch π'_{i+1} zur Gruppe G_{i+1} und ist also eine Zahl in K_{i+1}. Die Primzahl π'_{i+1} genügt aber offenbar einer Kongruenz

(34) $$x^p - \frac{1}{r_{i+1}}\tau^{b_{i+1}}\pi_i'^{\,e_{i+1}+1}x^{r_{i+1}} - \pi_i'\beta_i' \equiv 0 \pmod{\mathsf{P}^a}$$

in K_i, wobei

(35) $$\pi_i' \equiv \pi_i + \omega\,\pi_i^{\mu_\nu} \pmod{\mathsf{P}_i^{\mu_\nu+1}}$$

eine der Zahlen in (31) ist, und β_i' geht aus β_i hervor, wenn man π_i durch π_i' ersetzt.

Unter Anwendung der Resultate in A, Kap. 2, § 3 kann man aber zeigen, daß die Kongruenz (34) in K_{i+1} nicht lösbar sein kann. Wird zunächst nach (35)

$$\pi_i' = \pi_i \gamma_i = \pi_i(1 + A\pi_i^{\mu_\nu - 1}), \quad A \not\equiv 0 \pmod{\mathsf{P}_i}$$

gesetzt, so erhält die Kongruenz (34) die Form

$$x^p - \frac{1}{r_{i+1}} \tau^{b_{i+1}} \gamma_i^{e_{i+1}+1} \pi_i^{e_{i+1}+1} x^{r_{i+1}} - \pi_i \beta_i'' \equiv 0 \pmod{\mathsf{P}^a}, \tag{36}$$

wo

$$\beta_i'' \equiv \beta_i + A\pi_i^{\mu_\nu - 1} \pmod{\mathsf{P}_i^{\mu_\nu}}. \tag{37}$$

Nun ist es aber immer möglich, eine Zahl $\gamma_i' = 1 + B\pi_i^{\mu_\nu - 1}$ in K_i so zu bestimmen, daß

$$\gamma_i^{e_{i+1}+1} \cdot \gamma_i'^{\,p - r_{i+1}} \equiv 1 \pmod{\mathsf{P}_i^a},$$

und wenn man daher (36) mit $\gamma_i'^{\,p}$ multipliziert und x an der Stelle von $\gamma_i' x$ schreibt, so folgt, daß auch die Kongruenz

$$x^p - \frac{1}{r_{i+1}} \tau^{b_{i+1}} \pi_i^{e_{i+1}+1} x^{r_{i+1}} - \pi_i \beta_i''' \equiv 0 \pmod{\mathsf{P}^a} \tag{38}$$

gleichzeitig mit (34) in K_{i+1} lösbar sein muß. Dabei ist also nach (37)

$$\beta_i''' = \beta_i'' \gamma_i'^{\,p} \equiv \beta_i + A\pi_i^{\mu_\nu - 1} \pmod{\mathsf{P}_i^{\mu_\nu}}. \tag{39}$$

Wendet man aber auf (38) den Satz 4 in A, Kap. 2 an, so folgt sofort nach (39), daß, wenn die Kongruenz (38) in K_{i+1} lösbar sein soll, die Bedingung

$$\mu_\nu - 1 + r_{i+1} \equiv 0 \pmod{p} \tag{40}$$

erfüllt sein muß. Wenn aber π_i einer binomischen Kongruenz genügt, so ist $\varrho_i \equiv 0 \pmod{p}$, und folglich nach Satz 5 $\mu_\nu \equiv 1 \pmod{p}$, so daß nach (40) auch r_{i+1} durch p teilbar wäre, was offenbar nicht möglich ist. Es ist daher bewiesen, daß, wenn π_i einer binomischen Kongruenz genügt, die Kongruenz für π_{i+1} und daher für alle folgenden Primzahlen binomisch ist.

Es bleibt folglich nur übrig zu zeigen, daß, wenn π_i einer trinomischen Kongruenz genügt, dann π_{i+1} keiner binomischen Kongruenz genügen kann. Um dies zu leisten, wendet man auf π_{i+1} die Substitution S_i an, und wenn π_{i+1} einer binomischen Kongruenz (32) genügte, würde $\pi_{i+1}' = S_i : \pi_{i+1}$ der Kongruenz

$$x^p - \pi_i' \beta_i' \equiv 0 \pmod{\mathsf{P}^a} \tag{41}$$

genügen, wo die Bezeichnung die frühere ist. Die Kongruenz (41) kann dann weiter in der Form

$$(42) \qquad x^p - \pi_i \beta_i'' \equiv 0 \pmod{\mathsf{P}^a}$$

geschrieben werden, wo wie früher

$$(43) \qquad \beta_i'' \equiv \beta_i + A\pi_i^{\mu_v - 1} \pmod{\mathsf{P}_i^{\mu_v}}.$$

Wenn aber die Kongruenz (42) in K_{i+1} lösbar sein soll, muß nach Satz 2, A, Kap. 2 die Zahl

$$\frac{\beta_i''}{\beta_i} \equiv 1 + A\pi_i^{\mu_v - 1} \pmod{\mathsf{P}_i^{\mu_v}}$$

eine p-te Potenz $(\operatorname{mod} \mathsf{P}_i^a)$ sein, und dies ist weiter, wie leicht aus (28) folgt, nur dann möglich, wenn $\mu_v - 1$ durch p teilbar ist. Nach Satz 5 folgt aber daraus weiter, daß ϱ_i durch p teilbar ist, was nicht möglich sein kann, wenn π_i einer trinomischen Kongruenz genügt.

Die Normalkongruenzen sind also alle entweder binomisch oder trinomisch, und es bleibt nur zu zeigen, daß im trinomischen Falle der Exponent r_i eine von i unabhängige Konstante ist. Es genügt offenbar zu zeigen, daß $r_i \equiv r_{i+1} \pmod{p}$.

Dies folgt aber sehr einfach nach der eben angewandten Methode. Man wendet im trinomischen Falle auf π_{i+1} die Substitution S_i an, und die so erhaltene Primzahl π_{i+1}' wird einer Kongruenz (34) genügen. Wird weiter die Bedingung für die Lösbarkeit dieser Kongruenz in K_{i+1} gesucht, erhält man die Relation (40). Nach Satz 5 ist aber $\mu_v - 1 \equiv -r_i \pmod{p}$, und daraus folgt in der Tat $r_i \equiv r_{i+1} \pmod{p}$, wie bewiesen werden sollte.

Wenn man Satz 7 mit den Sätzen 5 und 2 kombiniert, erhält man sofort den weiteren interessanten Satz über die Verzweigungen:

Satz 8. *Im binomischen Falle ist*

$$(44) \qquad \mu_1 \equiv \mu_2 \equiv \ldots \equiv \mu_k \equiv 1 \pmod{p},$$

während im trinomischen Falle

$$(45) \qquad \mu_1 \equiv \mu_2 \equiv \ldots \equiv \mu_k \equiv 1 - r \pmod{p}.$$

Dieser Satz enthält speziell das Resultat von Speiser*), daß alle Verzweigungen einander $(\operatorname{mod} p)$ kongruent sein müssen.

Zuletzt sei auch noch erwähnt, daß man aus (15) und (16) noch die weitere Relation

$$(46) \qquad c_1 \equiv c_2 \equiv \ldots \equiv c_s \equiv -r \pmod{p-1}$$

erhält.

---

*) Speiser, Satz 4, S. 183.

Kapitel 2.
Der binomische Fall.
§ 1.
Bestimmung der Verzweigungen.

Der binomische Fall soll nun eingehend studiert werden, und wie man sehen wird, sind hier die Verhältnisse besonders einfach und übersichtlich.

Aus A, Kap. 2, Satz 5 folgt schon, daß, wenn der binomische Fall eintreten soll, der Regularkörper K_0 die p-te Einheitswurzel (mod P^a) enthalten muß. Wenn aber die Primzahl π_0 in K_0 durch die binomische Normalkongruenz (vgl. A, Kap. 1, Satz 6)

(1) $$x^{e_0} + \tau^{a_0} p \equiv 0 \pmod{\mathsf{P}^a}$$

definiert ist, so muß, als notwendige und hinreichende Bedingung, daß K_0 eine p-te Einheitswurzel (mod P^a) enthält,

$$e_0 \equiv a_0 \equiv 0 \pmod{p-1}$$

sein.

Weiter folgt aus Satz 2, daß für den i-ten Irregularkörper

(2) $$\varrho_i = e_0 p^i \qquad (i = 1, 2, \ldots, s)$$

ist, und da diese alle verschieden sind, folgt nach Satz 5, daß die Reihe der Verzweigungskörper mit der Reihe der Irregularkörper zusammenfällt, indem allgemein

$$K_i = K_{V_{i+1}} \qquad (i = 1, 2, \ldots, s).$$

Nach Satz 5 folgt weiter aus (2), daß die Verzweigung μ_i durch

$$\mu_i = \frac{e_0 p^i}{p-1} + 1 \qquad (i = 1, 2, \ldots, s)$$

bestimmt ist.

Satz 9. *Im binomischen Falle erhält man den Galoisschen Körper K aus dem Regularkörper K_0 durch eine Reihe von s Verzweigungskörpern von den sukzessiven Relativgraden p. Die Verzweigung μ_i hat allgemein den Wert*

(3) $$\mu_i = \frac{e_0 p^i}{p-1} + 1 \qquad (i = 1, 2, \ldots, s),$$

und die Konstanten des Regularkörpers müssen der Bedingung

(4) $$e_0 \equiv a_0 \equiv 0 \pmod{p-1}$$

genügen.

Zuletzt sei bemerkt, daß in diesem Falle natürlich alle Zahlen s_i in (2), Kap. 1 den gemeinsamen Wert 1 haben, und dann zeigt der Satz 3, Kap. 1, daß die Differente des Galoisschen Körpers durch P^{Δ} teilbar ist, wo

$$\Delta = (s+1)e - 1.$$

§ 2.

Sätze über die Verzweigungsgruppen.

Zuerst soll nun bewiesen werden:

Satz 10. *Im binomischen Falle ist die Verzweigungsgruppe zyklisch.*

Dieser Satz wird unter Anwendung einer Methode von Herrn Fueter[7]) einfach in der folgenden Weise bewiesen:

Im allgemeinen Falle, wo also das System der Normalkongruenzen beliebig binomisch oder trinomisch sein darf, sei V_i eine Substitution der i-ten Verzweigungsgruppe, folglich

$$V_i : \pi = \pi + \pi^{\mu_i} A_1(\pi), \quad A_1(\pi) = a_0 + a_1 \pi + \dots.$$

Durch Wiederholung folgt sofort

$$V_i^2 : \pi = \pi + 2\pi^{\mu_i} A_1(\pi) + \pi^{2\mu_i - 1} A_2(\pi), \quad A_2(\pi) = \mu_i a_0^2 + \dots,$$

$$V_i^3 : \pi = \pi + 3\pi^{\mu_i} A_1(\pi) + 3\pi^{2\mu_i-1} A_2(\pi) + \pi^{3\mu_i-2} A_3(\pi),$$

$$A_3(\pi) = \mu_i (2\mu_i - 1) a_0^3 + \dots,$$

und daher im allgemeinen

$$(5) \qquad V_i^p : \pi = \pi + \binom{p}{1} \pi^{\mu_i} A_1(\pi) + \binom{p}{2} \pi^{2\mu_i-1} A_2(\pi) + \dots + \pi^{p\mu_i - p + 1} A_p(\pi),$$

wo

$$(6) \qquad A_p(\pi) = a_0^p \mu_i (2\mu_i - 1) \dots ((p-1)\mu_i - p + 2) + \dots.$$

Nach der Ungleichung (28), Kap. 1 wird nun, das letzte Glied in (5) für die Verzweigungsgruppe bestimmend, in der V_i^p liegt.

Nimmt man nun zuerst den binomischen Fall an, so genügt μ_i nach (3) der Bedingung

$$(7) \qquad p\mu_i - p + 1 = \mu_{i+1},$$

und da in diesem Falle $\mu_i \equiv 1 \pmod p$, wird $A_p(\pi)$ nicht durch $\mathfrak{P}$ teilbar, so daß $V_i^p = V_{i+1}$ eine Substitution der $(i+1)$-ten, aber keiner höheren Verzweigungsgruppe wird; der Satz 10 ist dadurch bewiesen.

Im trinomischen Falle ist $\mu_i \not\equiv 1 \pmod p$ und $A_p(\pi)$ ist folglich durch $\mathfrak{P}$ teilbar; V_i^p gehört daher zu einer Verzweigungsgruppe G_{V_j}, $j > i$, wofür

$$\mu_j > p(\mu_i - 1) + 1.$$

Da aber nach Satz 8 $\mu_j \equiv -r + 1 \pmod p$ sein muß, erhält man den folgenden Satz, der zuerst von Speiser[8]) in einer etwas anderen Form abgeleitet worden ist:

---

[7]) R. Fueter, Ein Satz über Iteration von Potenzreihen und seine zahlentheoretische Anwendung, Vierteljahrsschrift d. Naturforschenden Ges. Zürich 1917, S. 67–72.

[8]) Speiser, Satz 5′, S. 184.

Satz 11. *Eine Substitution V_i^p gehört im trinomischen Falle zu einer Verzweigungsgruppe G_{V_j}, für die*

$$\mu_j \geqq p\mu_i - r + 1. \tag{8}$$

§ 3.

Eindeutigkeit der Normalkongruenzen.

Ehe wir zu einer weiteren Untersuchung der Gruppen im binomischen Falle übergehen, sollen ein paar wichtige Hilfssätze über die binomischen Normalkongruenzen bewiesen werden.

Es sei

$$x^p - \pi_i \beta_i \equiv 0 \pmod{\mathsf{P}^a} \tag{9}$$

die definierende Kongruenz für den $(i+1)$-ten Irregularkörper K_{i+1}. Die Zahl β_i hat nach Satz 2 die Form

$$\beta_i = 1 + \tau^{a_1}\pi_i + \ldots + \tau^{a_\nu}\pi_i^\nu, \tag{10}$$

wo $\nu = \frac{e_0 p^{i+1}}{p-1}$ ist. Es kann nun gezeigt werden, daß die Primzahl π_{i+1} in K_{i+1} sogar so gewählt werden kann, daß in (10) keine Glieder $\tau^{a_j}\pi_i^j$ vorkommen, wo der Exponent j durch p teilbar ist.

Wenn nämlich

$$\tau^{a_{tp}}\pi_i^{tp}, \qquad tp < \nu = \frac{e_0 p^{i+1}}{p-1} \tag{11}$$

das erste Glied dieser Art in (10) ist, kann man

$$\pi_{i+1}' = \pi_{i+1}\,(1 + \omega\pi_i^t)$$

setzen, und erhält für π_{i+1}' die Kongruenz

$$x^p - \pi_i\beta_i' \equiv 0 \pmod{\mathsf{P}^a},$$

wo β_i' den Wert

$$\beta_i' = \beta_i(1 + \omega\pi_i^t)^p = \beta_i + \beta_i\omega^p\pi_i^{tp} + \ldots + p\beta_i\omega\pi_i^t + \ldots \tag{12}$$

hat. Nach der Ungleichung (11) sind aber in (12) alle durch p teilbaren Glieder durch höhere Potenzen von P als die Zahl π_i^{tp} teilbar, so daß man für β_i' die Kongruenz

$$\beta_i' \equiv \beta_i + \omega^p\pi_i^{tp} \pmod{\mathsf{P}_i^{tp+1}}$$

erhält. Man braucht daher nur ω so zu bestimmen, daß

$$\omega^p + \tau^{a_{tp}} \equiv 0 \pmod{\mathsf{P}_i}$$

ist, was offenbar immer möglich ist. In dieser Weise kann man nach und nach alle Glieder (11) in (10) wegschaffen, indem diejenigen Glieder in β_i', welche höhere Potenzen als π_i^ν enthalten, nach A, Kap. 2, § 2 für die Lösbarkeit der Kongruenz keine Rolle spielen. Das letzte Glied in (10) ge-

hört auch zu den Ausnahmegliedern (11), aber es ist im allgemeinen nicht möglich dieses Glied wegzubringen.

Wenn β_i die Form hat, wo alle Glieder (11) fehlen, soll dies eine *reduzierte Darstellung* für β_i heißen. Die Bedeutung der reduzierten Darstellung geht aus dem folgenden Satze hervor:

Satz 12. *Es seien*

$$x^p - \pi_i \beta_i \equiv 0, \qquad x^p - \pi_i \beta_i' \equiv 0 \pmod{\mathsf{P}^a} \tag{13}$$

zwei Normalkongruenzen, welche beide K_{i+1} *definieren, und worin sowohl* β_i *als* β_i' *reduziert sind. Dann ist*

$$\beta_i' \equiv \beta_i \pmod{\mathsf{P}_i^\nu} \tag{14}$$

und weiter muß es eine Zahl ω *geben, so daß*

$$\tau^{a'_\nu} \equiv \tau^{a_\nu} + \omega^p - \omega\tau^{-a_0} \pmod{\mathsf{P}_i}. \tag{15}$$

Der Satz sagt kurz, daß die reduzierte Darstellung von β_i, abgesehen vom letzten Gliede, eine eindeutige ist.

Der Beweis ist einfach. Wenn die Kongruenzen (13) gleichzeitig in K_{i-1} lösbar sind, gibt es nach A, Kap. 2, Satz 2 eine solche Zahl γ_i in K_i, daß

$$\beta_i' \equiv \gamma_i^p \beta_i \pmod{\mathsf{P}^a}, \tag{16}$$

wo man natürlich γ_i in der Form

$$\gamma_i = 1 + A\pi_i^\sigma, \qquad A \not\equiv 0 \pmod{\mathsf{P}_i}. \tag{17}$$

schreiben kann. Für den Beweis des Satzes 12 ist offenbar nur der Fall von Interesse, daß $\sigma \leqq \frac{e_0 p^i}{p-1}$ ist. Wenn $\sigma < \frac{e_0 p^i}{p-1}$ ist, erhält man aus (16)

$$\beta_i' \equiv \beta_i + \beta_i A^p \pi_i^{\sigma p} + \ldots + \beta_i p A \pi_i^\sigma + \ldots \pmod{\mathsf{P}_i^a}. \tag{18}$$

Hier sind die durch p teilbaren Glieder durch höhere Potenzen von P_i als $\pi_i^{\sigma p}$ teilbar und folglich kann β_i' in diesem Falle nicht reduziert sein.

Man hat daher in (17) $\sigma = \frac{e_0 p^i}{p-1}$, und in diesem Falle ist nach (18) die Bedingung (14) erfüllt. Aus (18) erhält man dann aber weiter

$$\pi_i^\nu \tau^{a'_\nu} \equiv \pi_i^\nu \tau^{a_\nu} + A^p \pi_i^\nu + pA\pi^{\frac{e_0 p^i}{p-1}} \pmod{\mathsf{P}_i^{\nu+1}},$$

und wenn man hier den Wert von p aus (1) einsetzt, schließt man weiter

$$\tau^{a'_\nu} \equiv \tau^{a_\nu} + A^p - A\tau^{-a_0} \pmod{\mathsf{P}_i},$$

wie bewiesen werden sollte.

Man beweist auch leicht die Umkehrung des Satzes 12, daß, wenn (14) und (15) erfüllt sind, die Kongruenzen (13) gleichzeitig in K_{i+1} lösbar sein müssen.

Man kann in (15) voraussetzen, daß $\omega = \omega(\tau)$ eine Zahl im Trägheitskörper ist. Alle Zahlen

$$\omega^p - \omega\tau^{-a_0} \pmod{p}$$

in K_T bilden einen Modul

$$\mathsf{M} = (p, \omega^p - \omega\tau^{-a_0})$$

und man hat daher nach (15)

$$\tau^{a'_\nu} \equiv \tau^{a_\nu} \pmod{\mathsf{M}}.$$

Es folgt leicht, daß alle Zahlen im Trägheitskörper (mod M) in p Klassen von p^{f-1} Zahlen zerfallen. Denn eine Kongruenz

$$\omega_1^p - \omega\tau^{-a_0} \equiv \omega^p - \omega\tau^{-a_0} \pmod{p}$$

kann nur dann bestehen, wenn

$$(\omega_1 - \omega)^p \equiv (\omega_1 - \omega)\tau^{-a_0} \pmod{p},$$

woraus man sofort

$$\omega_1 \equiv \omega + k\tau^{\frac{a_0}{p-1}} \pmod{p} \quad (k = 0, 1, 2, \ldots, p-1)$$

erhält.

§ 4.

Reduktion der Normalkongruenzen.

Zuletzt soll noch gezeigt werden, wie man durch den Satz 12 eine noch weitere Reduktion der Normalkongruenzen erreichen kann.

Als Beispiel soll zuerst die erste Normalkongruenz

$$x^p - \beta_0\pi_0 \equiv 0 \pmod{\mathsf{P}^a}, \tag{19}$$

welche den ersten Irregularkörper definiert, studiert werden. Wie früher kann man annehmen, daß β_0 in reduzierter Darstellung gegeben ist:

$$\beta_0 = 1 + \tau^{a_1}\pi_0 + \ldots + \tau^{a_{\frac{e_0 p}{p-1}}}\pi_0^{\frac{e_0 p}{p-1}}, \tag{20}$$

worin, vom letzten Gliede abgesehen, keine Exponenten von π_0 durch p teilbar sind.

Wendet man nun auf die Wurzel π_1 der Kongruenz (19) eine Substitution T der Trägheitsgruppe an, so geht π_1 in eine Wurzel π'_1 der Kongruenz

$$x^p - \beta'_0\pi_0\tau^{b_0} \equiv 0 \pmod{\mathsf{P}^a}, \quad b_0 = \frac{p^f-1}{e_0} \tag{21}$$

über, wo also nach (20)

$$\beta'_0 = 1 + \tau^{a_1+b_0}\pi_0 + \tau^{a_2+2b_0}\pi_0^2 + \ldots. \tag{22}$$

Die Zahl $\pi_1' \cdot \tau^{\frac{b_0}{p}}$ genügt aber nach (21) der Kongruenz

$$x^p - \beta_0' \pi_0 \equiv 0 \pmod{\mathsf{P}^a}, \tag{23}$$

und diese Kongruenz muß gleichzeitig mit (19) in K_1 lösbar sein. Da aber nach (22) auch β_0' in reduzierter Form ist, so muß β_0', abgesehen vom letzten Gliede, mit β_0 identisch sein, und man hat daher für alle i

$$a_1 \equiv a_1 + i\, b_0 \pmod{p^f - 1},$$

d. h. i ist durch e_0 teilbar, und β_0 hat die einfache Form

$$\beta_0 = 1 + \tau^{a_1} \pi_0^{e_0} + \tau^{a_2} \pi_0^{\frac{e_0 p}{p-1}}. \tag{24}$$

Übt man weiter auf π_1 die Substitution Z aus der Zerlegungsgruppe aus, so geht π_0 in $\pi_0 \tau^{\lambda_0}$ über, und man erhält daraus sofort

$$a_1 \equiv p\, a_1 + e_0 \lambda_0 \pmod{p^f - 1},$$

und wenn man hier den Wert $\lambda_0 = \frac{a_0(p-1)}{e_0}$ einsetzt, kommt ohne Schwierigkeit

$$\tau^{a_1} \equiv k_0 \tau^{-a_0} \pmod{p},$$

wo k_0 eine rationale, nicht durch p teilbare Zahl ist. Nach (24) und (1) erhält daher β_0 die noch einfachere Form

$$\beta_0 = 1 + k_0 p + \tau^{a_2} \pi_0^{\frac{e_0 p}{p-1}}. \tag{25}$$

Im allgemeinen wird aber auch wegen der Bedingung (15) das letzte Glied in (25) verschwinden. Durch die Substitution T^i geht nämlich dieses Glied in

$$\tau^{a_2 + i b_0 \frac{e_0 p}{p-1}} \pi_0^{\frac{e_0 p}{p-1}}$$

über, und hier ist

$$\tau^{a_2 + i b_0 \frac{e_0 p}{p-1}} = \tau^{a_2 + i p \frac{p^f - 1}{p-1}} \equiv k_i \tau^{a_2} \pmod{p},$$

wo die rationale Zahl k_i alle Werte $1, 2, \ldots, p-1$ annehmen kann. Da nun nach Satz 12 immer eine Kongruenz

$$k_i \tau^{a_2} \equiv \tau^{a_2} + \omega^p - \tau^{-a_0} \omega \pmod{p}$$

bestehen muß, so folgt, wenn man $p > 2$ voraussetzt und daher $k_i = 2$ wählen kann,

$$\tau^{a_2} \equiv \omega^p - \tau^{-a_0} \omega \pmod{p},$$

d. h. τ^{a_2} gehört $(\mathrm{mod}\, p, \omega^p - \tau^{-a_0} \omega)$ zur selben Klasse wie die Zahl 0, und nach § 3 kann man dann die Zahl π_1 so wählen, daß das letzte

Glied in β_0 (25) nicht vorkommt. Im Falle $p=2$ kann man nach (25) $\beta_0 = 1 + 2k_0 + 4\omega_1$ schreiben, und wie früher zeigt man, daß die p^f möglichen ω_1 in zwei Klassen (mod 2, $\omega^2 - \omega$) zerfallen. Wenn f nicht durch 2 teilbar ist, gehören die Zahlen 0 und 1 zu verschiedenen Klassen, indem eine Kongruenz

$$\omega^2 - \omega \equiv 1 \pmod p$$

nicht lösbar sein kann. Man kann daher in diesem Falle $\omega_1 = 0$ oder $\omega_1 = 1$ annehmen. Wenn f gerade ist, wird man aber nicht immer das letzte Glied als rational annehmen können.

Satz 13. *Im binomischen Falle kann man die Primzahl π_1 des ersten Irregularkörpers so wählen, daß die Konstante β_0 durch*

$$\beta_0 = 1 + k_0 p \tag{26}$$

gegeben ist. Eine Ausnahme bildet nur der Fall, wo $p=2$ und f gerade ist; in diesem Falle hat man auch die Möglichkeit

$$\beta_0 = 1 + 2k_0 + 4\tau^{h_0}. \tag{27}$$

Im Falle $p > 2$ (und wir beschränken uns vorläufig nur auf diesen Fall) wird β_0 durch die Substitutionen T und Z nicht geändert, und man kann daher diese Substitutionen so wählen, daß

$$T: \pi_1 \equiv \tau^{\frac{h_0}{p}} \pi_1, \quad Z: \pi_1 \equiv \tau^{\frac{\lambda_0}{p}} \pi_1 \pmod{\mathsf{P}^a}.$$

Man kann nun durch Induktion für eine beliebige Normalkongruenz den folgenden Satz beweisen:

Satz 14. *Die Konstante β_i der $(i+1)$-ten Normalkongruenz hat für $p > 2$ die Form*

$$\begin{gathered} \beta_i = 1 + k_1 A_i + k_2 A_i^2 + \ldots, \\ A_i = \tau^{-\frac{a_0}{p^i}} \pi_i^{e_0}, \end{gathered} \tag{28}$$

wo alle Koeffizienten k rational sind; speziell fehlt das letzte Glied, welches die Potenz $\pi_i^{\frac{e_0 p^{i+1}}{p-1}}$ enthalten sollte. Weiter ist

$$T: \pi_{i+1} \equiv \tau^{\frac{h_0}{p^{i+1}}} \pi_{i+1}, \quad Z: \pi_{i+1} \equiv \tau^{\frac{\lambda_0}{p^{i+1}}} \pi_{i+1} \pmod{\mathsf{P}^a}. \tag{29}$$

Zunächst sei

$$\beta_i = 1 + \sum \tau^{a_r} \pi_i^r \tag{30}$$

die reduzierte Darstellung der Konstanten der $(i+1)$-ten Normalkongruenz

$$x^p - \beta_i \pi_i \equiv 0 \pmod{\mathsf{P}^a}. \tag{31}$$

Nimmt man nun den Satz 14 für alle vorangehenden Normalkongruenzen als bewiesen an, so folgt wie im Beweise des Satzes 13, wenn man auf (31) die Substitutionen T und Z anwendet, daß β_i in (30) die Form (28) haben muß, wenn man vom letzten Gliede absieht. Durch denselben Kunstgriff, wodurch das letzte Glied in β_0 weggebracht wurde, zeigt man aber, daß auch in β_i dieses Glied bei einer passenden Wahl von π_{i+1} zum Fehlen gebracht werden kann. Dann bleibt aber β_i durch Z und T ungeändert, und es folgt sofort, daß diese Substitutionen so bestimmt werden können, daß die Kongruenzen (29) bestehen. Der Satz 14 ist dadurch vollständig bewiesen.

Zuletzt sei noch erwähnt, daß man eine noch größere Reduktion der Koeffizienten β_i der Normalkongruenzen (31) erhält, wenn man noch die Bedingung ausnützt, daß diese Kongruenzen auch dann lösbar sein müssen, wenn man auf $\pi_i \beta_i$ eine Substitution V_j, $j < i$, einer der vorausgehenden Verzweigungsgruppen ausübt.

Aus diesen Betrachtungen erhält man den folgenden Satz, den ich ohne Beweis mitteile:

Satz 15. *Man kann immer die Konstante β_i der $(i+1)$-ten Normalkongruenz für $p > 2$ auf die Form*

$$(32)\qquad \begin{aligned} \beta_i &= 1 - \frac{1}{e_0(p^i-1)} A_i^{p^i-1} + p(l_1 A_i + l_2 A_i^2 + \ldots), \\ A_i &= \tau^{-\frac{a_0}{p^i}} \pi_i^{e_0} \end{aligned}$$

reduzieren, wo alle Koeffizienten l rational sind.

Aus (32) erhält man z. B. speziell für die Konstante der zweiten Normalkongruenz

$$\beta_1 = 1 - \frac{1}{e_0(p-1)} \tau^{-\frac{a_0}{p}(p-1)} \pi_1^{e_0(p-1)} + k p \tau^{-\frac{a_0}{p}} \pi_1^{e_0},$$

wo k eine beliebige rationale Zahl ist.

Wie man sieht, gibt der Satz (32) eine erhebliche Reduktion der Konstanten β_i. Die Form (32) repräsentiert aber im allgemeinen nicht die möglichst große Vereinfachung, und es ist wirklich möglich, eine absolut einfachste Normalform anzugeben. Es wird aber hier zu weit führen, diese Resultate abzuleiten. Sie sind aber, wie ich später erwähne, für die Bestimmung der Trägheitsgruppe in speziellen Fällen von Wichtigkeit.

§ 5.

Bestimmung der Trägheits- und Zerlegungsgruppe.

Nach diesen Vorbereitungen ist es im binomischen Falle möglich die vollständige Struktur der Zerlegungs- und Trägheitsgruppe anzugeben.

Da nach Satz 10 die Verzweigungsgruppe im binomischen Falle zyklisch ist, kann man eine solche Substitution V bestimmen, daß

$$G_V = V^k \qquad (k = 0, 1, \ldots, p^s - 1).$$

Weiter ist aber nach (29) für $i = s - 1$

$$(33) \qquad T : \pi \equiv \tau^{\frac{h_0}{p^s}} \pi, \quad Z : \pi \equiv \tau^{\frac{\lambda_0}{p^s}} \pi \pmod{\mathsf{P}^a}$$

und daraus folgt ohne Schwierigkeiten

$$T^{e_0} = 1, \quad Z^f = T^{a_0}, \quad Z^{-1} T Z = T^p.$$

Da G_V ein Normalteiler von G_Z und G_T ist, so wird

$$(34) \qquad Z^{-1} V Z = V^{z_0}, \quad T^{-1} V T = V^{t_0}$$

und unsere Aufgabe ist vollständig gelöst, wenn die Konstanten t_0 und z_0 $(\mathrm{mod}\, p^s)$ bestimmt sind.

Man kann zunächst die Eigenschaften von z_0 und $t_0 \pmod{p}$ ableiten. Wenn nämlich

$$V : \pi_1 = \pi_1 + \tau^{-\frac{a_1}{p-1}} \pi_1^{\frac{e_0 p}{p-1} + 1} + \ldots,$$

erhält man wie in (10), Kap. 1

$$T^{-1} V T : \pi_1 = \pi_1 + \tau^{-\frac{a_0}{p-1} + \frac{p^f - 1}{p-1}} \pi_1^{\frac{e_0 p}{p-1} + 1} + \ldots .$$

Hier ist aber

$$d_0 \equiv \tau^{\frac{p^f - 1}{p-1}} \pmod{p}$$

eine rationale Zahl, und da τ eine primitive Wurzel der Kongruenz

$$x^{p^f - 1} - 1 \equiv 0 \pmod{p^a}$$

ist, so folgt, daß $t_0 \equiv d_0 \pmod{p}$ eine primitive Wurzel der Kongruenz $x^{p-1} \equiv 1 \pmod{p}$ ist, d. h. t_0 gehört zum Exponenten $p - 1 \pmod{p}$. Man sieht auch leicht, daß durch eine passende Wahl von T die Zahl t_0 eine beliebige der $\varphi(p-1)$ zu $p - 1 \pmod{p}$ gehörenden Zahlen sein kann.

Aus (11), Kap. 1 erhält man in derselben Weise $z_0 \equiv 1 \pmod{p}$.

Um aber die Konstante t_0 vollständig zu bestimmen, bemerkt man, daß nach (34)

$$(35) \qquad T^{-(p-1)} V T^{p-1} = V^{t_0^{p-1}}.$$

Wenn nun

$$(36) \qquad \pi' = V : \pi = \pi + \sum_i \tau^{a_i} \pi^{r_i}$$

gesetzt wird, erhält man einfach nach (33)

$$\pi'' = T^{-(p-1)} V T^{p-1} : \pi = \pi + \sum_i \tau^{a_i + (r_i - 1)(p-1)\frac{h_0}{p^s}},$$

und folglich ist

$$(37) \qquad \pi' - \pi'' = \sum_i \tau^{a_i} \pi^{r_i} \left(1 - \tau^{(r_i - 1)(p-1)\frac{\mathfrak{b}_i}{p^s}}\right).$$

Diese Differenz muß aber natürlich genau durch $\mathsf{P}^{\frac{e_0 p^j}{p-1}+1}$ teilbar sein, wo j eine der Zahlen $1, 2, \ldots, s$ bezeichnet. Wird aber in (37) $r_i = \frac{e_0 p^j}{p-1} + 1$ gesetzt, so verschwindet der entsprechende Koeffizient in dieser Summe, und man hat folglich $\pi' = \pi''$. Nach (35) ist daher

$$T^{-(p-1)} V T^{p-1} = V,$$

und t_0 muß eine primitive Wurzel der Kongruenz

$$t_0^{p-1} \equiv 1 \pmod{p^s}$$

sein. Die Trägheitsgruppe ist dadurch vollständig bestimmt.

Satz 16. *Im binomischen Falle hat die Trägheitsgruppe die Form*

$$G_T = T^i V^j \quad (i = 0, 1, \ldots, e_0 - 1;\ j = 0, 1, \ldots, p^s),$$

wo

$$T^{e_0} = V^{p^s} = 1, \qquad T^{-1} V T = V^{t_0},$$

und t_0 ist eine primitive Wurzel der Kongruenz

$$t_0^{p-1} - 1 \equiv 0 \pmod{p^s}.$$

Dieser Satz ist nur für $p > 2$ bewiesen; durch einige einfache Übertragungen auf den Fall $p = 2$ zeigt man aber, daß der Satz auch für diese Primzahl richtig bleibt.

Aus Satz 13 folgt nämlich, daß β_0 durch die Substitution T nicht geändert wird, während π_0 in $\pi_0 \tau^{b_0}$ übergeht. Man kann daher T so wählen, daß $T: \pi_1 = \tau^{\frac{b_0}{2}} \pi_1$, und durch Induktion beweist man wie in (28), daß

$$\beta_i = 1 + \tau^{a_1} \pi_i^{e_0} + \tau^{a_3} \pi_i^{3 e_0} + \ldots + \tau^{a_{2^{i+1}}} \pi_i^{2^{i+1} e_0},$$

wo, abgesehen vom letzten Gliede, nur ungerade Exponenten vorkommen. β_i wird also auch durch T nicht geändert, und man erhält in dieser Weise bei einer passenden Wahl von T

$$T: \pi \equiv \tau^{\frac{b_0}{2^s}} \pi \pmod{\mathsf{P}^a},$$

woraus wie früher die Richtigkeit des Satzes 16 für $p = 2$ folgt. Es sei nebenbei bemerkt, daß die Trägheitsgruppe im Falle $p = 2$ Abelsch wird, und zwar zyklisch, indem die Exponenten e_0 und p^s relativ prim sind. Dies ist der einzige Fall, wo die Trägheitsgruppe bei binomischen Kongruenzen Abelsch werden kann.

Zuletzt soll noch die vollständige Zerlegungsgruppe bestimmt werden, und man braucht dafür nur die Konstante z_0 in (34) abzuleiten. Nach (34) folgt aber

$$Z^{-f} V Z^f = V^{z_0^f}$$

oder da $Z^f = T^{a_0}$, folgt aus Satz 16

$$T^{-a_0} V T^{a_0} = V^{t_0^{a_0}} = V,$$

indem a_0 nach Satz 9 durch $p-1$ teilbar ist. Man hat also

$$z_0^f \equiv 1 \pmod{p^s}.$$

Da aber $z_0 \equiv 1 \pmod{p}$, zeigt man leicht, wenn f nicht durch p teilbar ist, daß $z_0 \equiv 1 \pmod{p^s}$ sein muß. Wenn aber f genau durch p^{φ_0} teilbar ist, kann man nur $z_0 \equiv 1 \pmod{p^{s-\varphi_0}}$ schließen.

Satz 17. *Im binomischen Falle $p > 2$ hat die volle Zerlegungsgruppe die Form*

$$G_Z = Z^h T^i V^j \qquad \begin{matrix} (h = 0, 1, \ldots, f-1; \; i = 0, 1, \ldots, e_0 - 1) \\ (j = 0, 1, \ldots, p^s - 1), \end{matrix}$$

wo

$$Z^f = T^{a_0}, \qquad T^{e_0} = V^{p^s} = 1, \qquad Z^{-1} T Z = T^p$$

und

$$(38) \qquad Z^{-1} V Z = V^{z_0}, \qquad T^{-1} V T = V^{t_0},$$

wo t_0 eine beliebige primitive Wurzel der Kongruenz $t_0^{p-1} - 1 \equiv 0 \pmod{p^s}$ bezeichnet, und $z_0 \equiv 1 \pmod{p^{s-\varphi_0}}$, wenn f genau durch p^{φ_0} teilbar ist. Wenn f nicht durch p teilbar ist, kann man in (38) *einfach $z_0 = 1$ setzen.*

Es sei auch erwähnt, daß man wirklich Beispiele angeben kann, wo z_0 in (38) nicht gleich 1 ist. Die vollständige Bestimmung von z_0 hängt mit der am Ende des § 4 erwähnten größtmöglichen Reduktion der Konstanten β_i zusammen, und ich werde bei einer späteren Gelegenheit auf die Lösung dieses Problems zurückkommen.

Ebenso wird es hier zu weit führen, noch den Ergänzungssatz zu Satz 17 für $p = 2$ ausführlich zu beweisen. Es soll nur angegeben werden, daß, wenn in diesem Falle f ungerade ist, die Zerlegungsgruppe genau die Form des Satzes 17 hat, wobei in (38) $z_0 = t_0 = 1$ ist.

(Eingegangen am 21. 12. 1928.)

Über den sogenannten Produktsatz der Dimensionstheorie.

Von

Witold Hurewicz in Amsterdam.

1. Sind $A_1, A_2, \ldots, A_n$ irgendwelche topologische Räume und bedeutet a_i einen beliebigen Punkt von A_i, so nennt man die Menge aller Komplexe $(a_1, a_2, \ldots, a_n)$, nachdem dieselbe in bekannter Weise mit Umgebungen versehen ist, das *topologische Produkt*[1]) (oder kürzer den Produktraum) der n Räume A_i. Wir verwenden hierfür die Bezeichnung

$$[A_1 \times A_2 \times \ldots \times A_n].$$

In diesem Sinne ist also der Euklidische Zahlenraum R_n das topologische Produkt aus n Strecken.

Sind (wie wir es im folgenden stets voraussetzen wollen) die Räume $A_1, A_2, \ldots, A_n$ *metrische* Räume, so kann ihr Produktraum ebenfalls metrisiert werden, indem als Abstand zweier Punkte $(a_1, a_2, \ldots, a_n)$ und $(a'_1, a'_2, \ldots, a'_n)$ etwa die Zahl

$$\text{(a)} \qquad \sqrt{\sum_{i=1}^{n} (\overline{a_i, a'_i})^2}$$

definiert wird, wo $(\overline{a_i, a'_i})$ den Abstand zwischen den Punkten a_i und a'_i in A_i bedeutet.

Bemerken wir noch, daß das topologische Produkt aus kompakten Räumen immer selbst kompakt ist.

2. Bezüglich der Eigenschaften des Produktraumes in ihrer Abhängigkeit von den Eigenschaften der Faktorenräume liegen bis jetzt nur sehr wenige Ergebnisse vor. Die wichtigste Frage dieses Problemkreises betrifft die *Dimension* des Produktraumes.

Naheliegend ist die Vermutung — sie wurde von Menger als der *Produktsatz der Dimensionstheorie* bezeichnet —, daß die Dimension des

[1]) Vgl. Tietze, Math. Annalen **88**, S. 298 (das topologische Produkt wird von Tietze auch für unendlich viele Faktoren definiert).

Produktraumes immer aus den Dimensionen der Faktoren durch Addition entsteht:

$$\dim [A_1 \times A_2] = \dim A_1 + \dim A_2 .$$

Da die Ungleichung: $\dim [A_1 \times A_2] \leqq \dim A_1 + \dim A_2$ schon mit den einfachsten Mitteln bewiesen werden kann[2]), reduziert sich der wesentliche Inhalt des Produktsatzes auf die Behauptung:

$$\dim [A_1 \times A_2] \geqq \dim A_1 + \dim A_2 .$$

Ein Spezialfall des Produktsatzes (und zwar, von dem trivialen Fall der nulldimensionalen Faktoren abgesehen, der einzige Spezialfall, der bis jetzt erledigt wurde) ist der Brouwersche „Rechtfertigungssatz des Dimensionsbegriffes"[3]), welcher besagt, daß der R_n (das topologische Produkt aus n Strecken) die Dimension n hat.

Wenn der Produktsatz richtig ist, dann muß offenbar die folgende Aussage gelten: Das Produkt von n Räumen, deren jeder eine positive Dimension besitzt, ist mindestens n-dimensional. Im folgenden soll diese Behauptung unter Beschränkung auf metrische kompakte Räume bewiesen werden. Da jeder kompakte metrische Raum von einer positiven Dimension bekanntlich ein Teilkontinuum enthält[4]), können wir den ausgesprochenen Satz auch in die folgende Gestalt bringen:

(A) *Das Produkt von n kompakten (metrischen) Kontinua ist mindestens n-dimensional.*

Daraus folgt insbesondere, daß das Produkt von n eindimensionalen kompakten Räumen genau n-dimensional ist, was eine Verallgemeinerung des oben erwähnten Brouwerschen Theorems ist. Auf jenes Theorem wird die Behauptung (A) letzten Endes auch zurückgeführt.

3. Der Beweis des Theorems (A) erfordert einige vorbereitende Betrachtungen. Wir knüpfen an den bekannten dimensionstheoretischen Satz von Menger und Urysohn an:

I. Ist der kompakte Raum R höchstens $(n-1)$-dimensional, so kann zu jeder positiven Zahl ε ein System aus endlich vielen abgeschlossenen Teilmengen von R angegeben werden, die 1. in ihrer Gesamtheit den ganzen Raum R ausfüllen, die 2. durchwegs von Durchmessern $< \varepsilon$ sind und die 3. zu je $n+1$ fremd sind[5]).

---

[2]) Vgl. Menger, Dimensionstheorie (1928), S. 246.

[3]) Brouwer, Journ. f. Math. **142**, S. 148; vgl. Menger, a. a. O. S. 244; ferner s. unten [10]).

[4]) Vgl. Menger, a. a. O. S. 213.

[5]) Vgl. Menger, a. a. O. S. 155 ff. Nach Urysohn gilt auch die Umkehrung des Theorems (wegen des Beweises siehe etwa Menger, a. a. O. S. 174 ff.).

Ein System von abgeschlossenen Mengen, das nur der Bedingung 1 genügt, heißt eine *Überdeckung* des Raumes R; ist überdies noch die Bedingung 2 erfüllt, so sprechen wir von einer *ε-Überdeckung*. Von einer Überdeckung, welche die Bedingung 3 befriedigt, sagen wir im Anschluß an Urysohn, sie sei von einer *Ordnung* $\leqq n$, wobei also, genauer gesprochen, unter der Ordnung einer Überdeckung die größte natürliche Zahl m verstanden wird, für die es Punkte gibt, die m Mengen der Überdeckung gemeinsam sind.

Dem angeführten Satz zufolge ist das Nicht-Vorhandensein von ε-Überdeckungen von einer Ordnung $\leqq n$ bei genügend kleinem ε eine hinreichende Bedingung dafür, daß der kompakte Raum R mindestens n-dimensional sei. Um aus diesem *negativen* Kriterium ein *positives*[5a]) herzuleiten, führen wir die folgende Begriffsbildung ein:

Sind zwei Überdeckungen U und V desselben Raumes R gegeben, wobei jede Menge des Systems V entweder mit einer Menge des Systems U identisch ist, oder Summe von einigen Mengen des Systems U ist, so nennen wir V eine aus der Überdeckung U *abgeleitete Überdeckung*.

Wir formulieren nun den Satz:

II. *Damit der kompakte Raum R eine Dimension $\geqq n$ habe, ist hinreichend, daß es eine Zahl $\eta > 0$ von der folgenden Eigenschaft gebe: Zu jedem noch so kleinen positiven ε existiert eine ε-Überdeckung U_ε derart, daß sämtliche aus ihr abgeleitete η-Überdeckungen eine Ordnung $> n$ haben*[6])[7]).

Nehmen wir zum Beweis an, der Raum R, für den wir die Bedingung des Satzes als erfüllt voraussetzen, sei höchstens $(n-1)$-dimensional; dann gibt es nach Satz I eine η-Überdeckung von einer Ordnung $\leqq n$; diese Überdeckung möge aus den Mengen

$$A_1, A_2, \ldots, A_m$$

bestehen, die nach Annahme zu je $n+1$ fremd sind. Wir können eine positive Zahl ε bestimmen, derart, daß die Mengen $U(A_i, \varepsilon)$ (darunter wird üblicherweise die Menge aller Punkte verstanden, deren Abstand von der abgeschlossenen Menge A_i kleiner ist als ε) gleichfalls zu je $n+1$

5a) Vgl. in diesem Zusammenhange die Stellungnahme Alexandroffs (Göttinger Nachrichten, Juli 1928, S. 26), nach dessen Ansicht „alle bis jetzt bekannten Eigenschaften der *mindestens* n-dimensionalen Mengen nicht nur der Form, sondern auch dem Inhalt nach einen ausgesprochen negativen Charakter haben".

6) Die Bedingung ist auch notwendig, wie in trivialer Weise aus dem sub 5) zitierten Urysohnschen Satze hervorgeht.

7) Die Tatsache, daß die Voraussetzungen von Satz I und von Satz II äquivalent sind, wurde implizite von Lebesgue verwendet; vgl. Fund. Math. 2, S. 257.

fremd sind; dabei wählen wir ε so klein, daß die Durchmesser der $U(A_i, \varepsilon)$ unterhalb η verbleiben.

Sei jetzt U_ε eine gemäß der Bedingung des zu beweisenden Satzes gewählte ε-Überdeckung von R, und sei B_i $(i = 1, 2, \ldots, m)$ die Summe aller derjenigen Mengen des Systems U_ε, welche mit A_i gemeinsame Punkte haben. Es ist dann

$$(') \qquad A_i \subset B_i \subset U(A_i, \varepsilon).$$

Die m Mengen B_i bilden eine aus U_ε abgeleitete Überdeckung von R, und aus den Beziehungen (') geht hervor, daß dies eine η-Überdeckung ist und daß ihre Ordnung $\leqq n$ ist, was im Widerspruch steht mit der Wahl der Überdeckung U_ε. Es ist somit gezeigt, daß R mindestens n-dimensional ist.

4. Wir erinnern jetzt an die folgende einfache Eigenschaft der kompakten Kontinua:

Ist C ein kompaktes Kontinuum und ε eine beliebige positive Zahl, so kann eine endliche Folge von abgeschlossenen Teilmengen von C

$$(\mathrm{X}) \qquad M_1, M_2, \ldots, M_k$$

angegeben werden, so daß 1. die Mengen M_i eine ε-Überdeckung von C bilden und daß 2. je zwei aufeinanderfolgende Mengen M_i und M_{i+1} mindestens einen gemeinsamen Punkt haben.

Aus der Kompaktheit von C folgt nämlich zunächst die Existenz einer ε-Überdeckung, und aus dem Zusammenhang von C ergibt sich sodann, daß die Mengen dieser Überdeckung in der Weise angeordnet werden können (wobei sie auch mehrfach gezählt werden dürfen), daß die Bedingung 2. realisiert wird[8]).

Wir fügen diesem Satz die folgende Bemerkung hinzu: Sind p und q zwei beliebig gegebene Punkte von C, so kann die Folge (X) immer so gewählt werden, daß die erste Menge M_1 den Punkt p und die letzte M_k den Punkt q enthalte. Nehmen wir nämlich an, p komme in M_1 nicht vor, und sei $i > 1$ der erste Index, für den $p \subset M_i$, dann ersetzen wir die Folge (X) durch die Folge:

$$M_i, M_{i-1}, \ldots, M_2, M_1, M_2, \ldots, M_i, \ldots, M_k.$$

Durch eine analoge Modifikation erreicht man, daß der Punkt q der letzten Menge der Folge angehört.

5. Nun sind wir mit Hilfsmitteln genügend ausgerüstet, um an den Beweis des im § 2 formulierten Theorems (A) herangehen zu können.

Es seien also

$$C_1, C_2, \ldots, C_n$$

---

[8]) Siehe etwa v. Kerékjártó, Vorlesungen über Topologie, S. 100.

n kompakte Kontinua. Wir haben zu zeigen, daß

$$\dim[C_1 \times C_2 \times \ldots \times C_n] \geqq n$$

ist.

Bezeichnen wir mit η den kleinsten unter den Durchmessern der n Kontinua C_i (es ist natürlich $\eta > 0$) und wählen in jedem C_i ein für allemal je zwei Punkte a_i und b_i, deren Abstand nicht kleiner ist als η.

Sei ε eine willkürliche positive Zahl, die wir im folgenden festhalten wollen. In jedem C_i bestimmen wir eine endliche Folge von abgeschlossenen Teilen:

$$\text{(v)} \qquad M_1^i, M_2^i, \ldots, M_k^i \qquad (i = 1, 2, \ldots, n),$$

welche eine ε-Überdeckung von C_i bilden und so angeordnet sind, daß je zwei aufeinanderfolgende einen gemeinsamen Punkt haben und daß a_i in M_1^i und b_i in M_k^i liegt. Die Anzahl k der Mengen haben wir als dieselbe für alle C_i angenommen, was natürlich zulässig ist, da wir die Anzahl der Mengen in der Folge (v) durch mehrfache Zählung beliebig vergrößern können.

Betrachten wir jetzt den Produktraum:

$$P = [C_1 \times C_2 \times \ldots \times C_n]$$

und denken wir uns denselben metrisiert nach der Formel (a) in § 1. Die Mengen

$$\text{(v)} \qquad M_{i_1 i_2 \ldots i_n} = [M_{i_1}^1 \times M_{i_2}^2 \times \ldots \times M_{i_n}^n] \quad (i_1, i_2, \ldots, i_n = 1, 2, \ldots, k)$$

bilden eine Überdeckung von P, und ihre Durchmesser sind kleiner als $\varepsilon\sqrt{n}$ (also mit ε beliebig klein). Wir wollen zeigen, daß die so hergestellte Überdeckung von P in Bezug auf die Zahl η die im Satz II von § 3 angegebene Eigenschaft besitzt.

6. Wir greifen aus dem System (v) eine Anzahl r von Mengen heraus:

$$\text{(w)} \qquad M_{i_1^1 i_2^1 \ldots i_n^1}, M_{i_1^2 i_2^2 \ldots i_n^2}, \ldots, M_{i_1^r i_2^r \ldots i_n^r},$$

und fragen: unter welchen Bedingungen haben diese r Mengen einen gemeinsamen Punkt? Nennen wir zur Abkürzung zwei n-Tupel natürlicher Zahlen $(i_1, i_2, \ldots, i_n)$ und $(j_1, j_2, \ldots, j_n)$ *benachbart*, wenn jede der Differenzen $i_m - j_m$ $(m = 1, 2, \ldots, n)$ entweder $= 0$ oder $= 1$ ist. Dann gilt: Damit die Mengen (w) einen gemeinsamen Punkt haben, ist *hinreichend*, daß je zwei von den r Komplexen

$$(i_1^m, i_2^m, \ldots, i_n^m) \qquad (m = 1, 2, \ldots, r)$$

benachbart seien.

Ist nämlich diese Voraussetzung erfüllt, so gibt es bei festem $l \leqq n$ unter den r Indizes

$$i_l^1, i_l^2, \ldots, i_l^r$$

höchstens zwei verschiedene, die sich dann um eine Einheit unterscheiden. Das System der Teilmengen

$$M_{i_l^1}^l, M_{i_l^2}^l, \ldots, M_{i_l^r}^l$$

von C_l reduziert sich demnach entweder auf eine einzige Menge oder auf zwei *nicht-fremde* Mengen. Jedenfalls gibt es also (für $l = 1, 2, \ldots, n$) in C_l einen Punkt — wir nennen ihn p_l —, so daß für $k = 1, 2, \ldots, r$ $p_l \subset M_{i_l^k}^l$. Der Punkt $(p_1, p_2, \ldots, p_n)$ von P gehört dann, wie leicht ersichtlich, jedem der Produkte

$$M_{i_1^m i_2^m \ldots i_n^m} = [M_{i_1^m}^1 \times M_{i_2^m}^2 \times \ldots \times M_{i_n^m}^n] \qquad (m = 1, 2, \ldots, r)$$

an.

7. Der weitere Gedankengang besteht darin, daß wir das System der Mengen $M_{i_1 i_2 \ldots i_n}$ hinsichtlich der Inzidenzbeziehungen zwischen diesen Mengen mit einem System von Intervallen des R_n vergleichen, dessen Eigenschaften uns wohlbekannt sind.

Sei W das Einheitsintervall des Koordinatenraumes R_n:

$$0 \leqq x_m \leqq 1 \qquad (m = 1, 2, \ldots, n).$$

Wir teilen W in k^n Würfel $W_{i_1 i_2 \ldots i_n}$ $(i_1 \ldots i_n = 1, 2, \ldots, k)$ von der Kantenlänge $\frac{1}{k}$ ein; $W_{i_1 i_2 \ldots i_n}$ ist durch die Koordinatenungleichungen:

$$\frac{i_m - 1}{k} \leqq x_m \leqq \frac{i_m}{k} \qquad (m = 1, 2, \ldots, n)$$

definiert. Eine elementargeometrische Betrachtung lehrt, daß r Würfel $(r \leqq k^n)$

$$\text{(ww)} \qquad W_{i_1^1 i_2^1 \ldots i_n^1}, W_{i_1^2 i_2^2 \ldots i_n^2}, \ldots, W_{i_1^r i_2^r \ldots i_n^r}$$

dann und nur dann einen gemeinsamen Punkt haben, wenn die entsprechenden r Indizeskomplexe zu je zwei benachbart sind. Die Bemerkung des vorigen Paragraphen können wir daher in die Gestalt bringen:

Wenn die r Würfel (ww) einen gemeinsamen Punkt haben, so auch die r Mengen (w).

8. Nunmehr bilden wir aus der Überdeckung (v) von P eine abgeleitete η-Überdeckung[8a]):

$$S_1, S_2, \ldots, S_m,$$

[8a]) η bedeutet dabei, wie im § 5, den größten unter den Durchmessern der C_l.

wo also S_i Summe ist von einigen Mengen $M_{i_1 \dots i_n}$ (bzw. mit einer dieser Mengen identisch); indem wir die *entsprechenden* (d. h. mit denselben Indizes versehenen) Würfel $W_{i_1 \dots i_n}$ in Würfelsummen Z_i vereinigen, erhalten wir eine Überdeckung

$$Z_1, Z_2, \dots, Z_m$$

von W. Wir beweisen die folgende Eigenschaft der letzteren Überdeckung:

Keiner der Würfelkomplexe Z_i besitzt Punkte auf gegenüberliegenden $(n-1)$-dimensionalen Seiten von W.

Angenommen nämlich, einer der genannten Würfelkomplexe, etwa Z_i, berühre ein Paar von gegenüberliegenden Seiten von W; der Einfachheit halber wollen wir voraussetzen, dies seien die Seiten $x_1 = 0$ und $x_1 = 1$. Unter den Würfeln, aus denen Z_i zusammengesetzt ist, kommen dann zwei

$$W_{i_1 i_2 \dots i_n}, \quad W_{j_1 j_2 \dots j_n}$$

vor mit

$$i_1 = 1, \quad j_1 = k.$$

Betrachten wir die entsprechenden Produktmengen

$$(') \quad \begin{aligned} M_{1, i_2, \dots, i_n} &= [M_1^1 \times M_{i_2}^2 \times \dots \times M_{i_n}^n], \\ M_{k, j_2, \dots, j_n} &= [M_k^1 \times M_{j_2}^2 \times \dots \times M_{j_n}^n], \end{aligned}$$

welche beide Teilmengen von S_i sind, und erinnern wir uns, daß es im Kontinuum C_1 zwei Punkte (nämlich die im § 5 verwendeten Punkte a_1 und b_1) mit einem Abstand $\geqq \eta$ gibt, deren ersterer in M_1^1, und deren zweiter in M_k^1 enthalten ist. Daraus geht hervor, daß man im Produktraum P zwei Punkte bestimmen kann, die resp. in den Mengen (') (also beide in S_i) gelegen sind und einen Abstand $\geqq \eta$ voneinander haben[9]). Dann ist also der Durchmesser von S_i nicht kleiner als η im Widerspruch mit der anfangs gemachten Voraussetzung, die Überdeckung $\{S_i\}$ sei eine η-Überdeckung.

Aus der bewiesenen „Randeigenschaft" der Mengen Z_i folgt nach einem fundamentalen Satze von Lebesgue und Brouwer[10]), daß es in W mindestens

[9]) Ist nämlich q_i ein beliebig gewählter Punkt von C_i, so haben die Punkte $(a_1, q_{i_2}, q_{i_3}, \dots, q_{i_n})$ und $(b_1, q_{j_2}, q_{j_3}, \dots, q_{j_n})$ des Produktraumes P die erwähnte Eigenschaft.

[10]) Vgl. Lebesgue, Math. Annalen **70**, S. 166 und Fund. Math. **2**, S. 257. Das erste Mal wurde das Theorem von Brouwer (Journ. f. Math. **142**, S. 149) bewiesen. Es bildet die Grundlage für den oben erwähnten Brouwerschen Rechtfertigungssatz des Dimensionsbegriffes, so daß also in der vorliegenden Arbeit, wie bereits oben hervorgehoben, der Nachweis der n-Dimensionalität von P im wesentlichen auf die bekannte Tatsache, daß R n-dimensional ist, zurückgeführt wird. Einfache Beweise des Lebesgue-Brouwerschen Theorems finden sich bei Sperner (Hamburger Berichte **6**, S. 265) und bei mir (Math. Annalen **101** (1929), S. 210).

einen Punkt gibt, der $n+1$ der Komplexe Z_i gemeinsam ist, daß m. a. W. $n+1$ Würfel

$$W_{i_1^1 i_2^1 \ldots i_n^1},\ W_{i_1^2 i_2^2 \ldots i_n^2},\ \ldots,\ W_{i_1^{n+1} i_2^{n+1} \ldots i_n^{n+1}}$$

existieren, die einen gemeinsamen Punkt besitzen, und dabei der Reihe nach $n+1$ verschiedenen Komplexen Z_i angehören. Gehen wir zu den entsprechenden Teilmengen von P

$$M_{i_1^1 i_2^1 \ldots i_n^1},\ M_{i_1^2 i_2^2 \ldots i_n^2},\ \ldots,\ M_{i_1^{n+1} i_2^{n+1} \ldots i_n^{n+1}}$$

über, so haben auch diese Mengen nach § 6 einen gemeinsamen Punkt und sind vermöge der Korrespondenz zwischen den S_i und den Z_i der Reihe nach in $n+1$ verschiedenen Mengen S_i enthalten. Somit gibt es Punkte von P, die $n+1$ der Mengen S_i gemeinsam sind, d. h. die Ordnung der Überdeckung $\{S_i\}$ ist mindestens $n+1$. Da $\{S_i\}$ eine beliebige aus der $\varepsilon\sqrt{n}$-Überdeckung $\{M_{i_1 \ldots i_n}\}$ abgeleitete η-Überdeckung war, so ist nach Satz II in § 3 der Raum P mindestens n-dimensional. Das am Anfang der Arbeit ausgesprochene Theorem ist somit bewiesen[11]).

[11]) Es ist zu vermuten, daß das Produkt von n Kontinua nicht nur n-dimensional, sondern sogar n-*stufig zusammenhängend* sein muß. (Wegen des letzten Begriffes vgl. Menger a. a. O., S. 214ff.)

(Eingegangen am 18. 10. 1928.)

Concerning points of continuous curves defined by certain im kleinen properties[1]).

Von

G. T. Whyburn in Austin (Texas, U. S. A.).

§ 1.

Introduction.

This paper contains primarily the results of a study of the *im kleinen cut points*[2]) and *im kleinen cycle points* of a continuous curve M in euclidean space of n dimensions. If for each $\varepsilon > 0$, a domain R exists containing the point P of M and of diameter $< \varepsilon$ such that P is a cut point of the component of $M \cdot \bar{R}$ which contains P, then P is said to be *im kleinen cut point* of M; if P lies, for each ε, on some simple closed curve in M of diameter $< \varepsilon$, P is said to be an *im kleinen cycle point* of M. § 2 contains various characterizations of these types of points together with demonstrations of certain of their properties. It is there shown that the set of all im kleinen cut points of any continuous curve is a Borel set of the class F_σ (*i. e.* the sum of a countable number of closed sets); the problem of determining the Borel class, if any, of the set of all im kleinen cycle points is left open. In § 3 the possibility of the density of the non-im kleinen cut points L and the ramification points (points

[1]) Presented to the American Mathematical Society, June 2, 1928.

[2]) The point P of a connected set M is said to be a cut point of M provided $M - P$ is not connected. The notion of an im kleinen cut point of a continuum is contained implicitly in the works of P. Urysohn and R. L. Moore, and is closely approximated in that of R. G. Lubben and C. Zarankiewicz. Cf. P. Urysohn, Über im kleinen zusammenhängende Kontinua, Math. Annalen **98** (1927), S. 296—308; [Urysohn uses the terms „unvermeidbar" (unavoidable) and „vermeidbar" (avoidable) to designate im kleinen cut points and non-im kleinen cut points, respectively]; R. L. Moore, Concerning Triods in the Plane and the Junction Points of Plane Continua, Proc. Ntl. Acad. of Sci. **14** (1928), pp. 85—88; R. G. Lubben, Concerning Connectedness near a Point Set; and C. Zarankiewicz, Sur les points de division dans les ensembles connexes, Fund. Math. **9** (1927), see proof of Theorem 14.

of Menger order > 2) W of M on an arc t of M is investigated and it is found that if L is dense on t it must be uncountably dense on t and if M is cyclicly connected and W is dense on t, then L must be uncountably dense on t. In § 4 a study is made of a continuous curve M composed wholly, or almost wholly, of im kleinen cut points. Some more general theorems are established from which it follows that if M is composed wholly of im kleinen cut points, then M is a Menger regular curve which, if bounded, is for each $\varepsilon > 0$, the sum of a finite number of ε-continua no two having more than one point in common; and if M is cyclicly connected and W denotes the set of all its ramification points, then $\overline{W}$ is totally disconnected and each component of $M - \overline{W}$ is an ordinary arc-segment (*i. e.*, a simple continuous arc minus its endpoints). A special type of regular curve, called a *node curve*, is studied in § 5. These curves are defined as continua M which, for each $\varepsilon > 0$, are the sum of a finite number of ε-continua each having at most two points in common with the rest of M. The results of § 5 show that this class of curves includes all acyclic continuous curves [*i. e.*, continuous curves containing no simple closed curve, or baum curves (Menger)], and all baum im kleinen curves. In § 6 it is shown that an im kleinen cut point of a continuous curve M may be characterized as a point which is an isolated point of some irreducible cutting of M between some pair of points A and B of M, and that in a Menger regular curve the set of all non-im kleinen cut points of M is totally disconnected. In § 7 an example in 3-space is constructed of a continuous curve every subcontinuum of which is a continuous curve which has a number of interesting properties, among them being that it contains infinitely many mutually exclusive arcs all of diameter $> 1/2$. In this section also is given theorems and discussion of the extension to n-space of some known theorems about Menger regular curves and continuous curves all of whose subcontinua are continuous curves in the plane.

The term continuous curve is used in this paper to designate any connected im kleinen continuum, bounded or not. The point sets considered are assumed to lie in a euclidian n-space, although it is obvious from the proofs that many of the theorems hold in more general spaces.

Definitions. A continuum will be called an ε-continuum, or in general, a set will be called an ε-set, provided that continuum or set is of diameter $< \varepsilon$, where ε denotes some positive number given in advance. A point P of continuum M is a regular point[3]) of M provided that for

[3]) *Cf.* K. Menger, Grundzüge einer Theorie der Kurven, Math. Annalen **95** (1925), S. 272–306.

each $\varepsilon > 0$, an ε-neighborhood R of P exists such that $F(R) \cdot M$ is finite, where as in this paper it will be used, $F(R)$ denotes the boundary of R. If an integer n exists such that for each ε, the neighborhood R can be chosen so that $F(R) \cdot M$ contains at most n points and if n is the smallest integer such that this property holds, then P is said to be a point of order[4]) n of M. Points of order 1 of continuum are called endpoints[5]), points of order > 2 are called ramification points[6]), and regular points which have no finite order are said to be of order w. A continuum all of whose points are regular will be called a Menger regular curve or simply a regular curve. A continuous curve M is said to cyclicly connected[7]) provided that every two points of M lie together on some simple closed curve in M. A cyclicly connected continuous curve C is called a maximal cyclic curve of a continuous curve M provided C is a subset of M but is not a proper subset of any cyclicly connected continuous curve in M. Considerable use is made in this paper of the decomposition of a continuous curve into its cyclic elements, *i. e.*, maximal cyclic curves, cut points, and endpoints, an extensive theory of which may be found in my paper *Concerning the Structure of a Continuous Curve*[8]).

The ordinary notation of Point Set Theory will be used in this paper. In general the letter M is used to denote a continuous curve or a regular curve, and, unless otherwise stated, the letters K, N, H and W denote the

---

[4]) *Cf.* K. Menger, *loc. cit.*, and P. Urysohn, Comptes Rendus **175** (1922), p. 481. Urysohn uses the term 'index of a point' instead of the therm 'order of a point'.

[5]) That this definition is equivalent for the case of continuous curves to the Wilder definition [R. L. Wilder, Concerning Continuous Curves, Fund. Math. **7** (1925), pp. 340–377] was shown by H. M. Gehman. See Concerning End Point of Continuous Curves and other Continua, Trans. Amer. Math. Soc. **30** (1928). In my thesis I showed that this definition is (for plane continuous curves M) equivalent to the following simple one: P is an endpoint of M provided that P is an interior point of no arc in M. *Cf.* Concerning Continua in the Plane, Trans. Amer. Math. Soc. **29** (1927), pp. 369–400, Theorem 12. For the extension of this and other results frequently used later to n-space see W. L. Ayres, Concerning Continuous Curves in a Space of n Dimensions, Amer. Journal of Math.

[6]) *Cf.* W. Sierpinski, Comptes Rendus **160**, p. 305. Sierpinski defines a ramification point of M as a point P such that M contains 3 continua K, L and N, such that $K \cdot L = K \cdot N = L \cdot N = P$. It follows from a result of Menger's (Fund. Math. **10**) that the definition here given and Sierpinski's definition are equivalent for Menger regular curves. Rutt [Bull. Amer. Math. Soc. **33** (1927), p. 411 (abstract)] has shown them equivalent for all plane continuous curves. It appears likely that they are equivalent for continuous curves in n-space.

[7]) *Cf.* my paper Cyclicly Connected Continuous Curves, Proc. Ntl. Acad. of Sci. **13** (1927), pp. 31–38.

[8]) Amer. Journ. of Math. **50** (1928), pp. 167–194.

sets of all im kleinen cut points, im kleinen cycle points, end points, and ramification points of the curve M. The symbol $S(P, r)$ denotes the set of all points whose distance from the point P is less than the number r.

§ 2.

Im Kleinen Cut Points and Im Kleinen Cycle Points.

Theorem 1. *In order that the point P of a continuous curve M should be an im kleinen cut point of M it is necessary and sufficient that P should be a cut point of some connected open*[9]) *subset of M.*

Theorem 2. *In order that the point P of a bounded continuous curve M should be a non-im kleinen cut point* (*or an avoidable point in the terminology of Urysohn*) *it is necessary and sufficient that $M - P$ should be uniformly connected im kleinen*[10]).

Theorem 3. *Suppose R_0 and R are bounded connected open subsets of a continuous curve M and suppose $\bar{R} \subset R_0$. Then there exists a continuous curve U such that* (1) $R \subset U \subset R_0$, (2) *every cut point of R_0 which belongs to R is a cut point of U, and* (3) *every cut point of U is a cut point of R and hence is an im kleinen cut point of M.*

The proofs of Theorems 1 and 2 present no difficulties. Theorem 3 is readily established with the aid of Theorem 1 above and Theorem 1 in the paper *On continuous curves in n dimensions*[11]) by W. L. Ayres and the author.

Theorem 4. *If the im kleinen cut point P of a continuous curve M is a point of order two of M, then P is not an im kleinen cycle point of M.*

Proof. Suppose, on the contrary, that P is an im kleinen cycle point of M. There exists a neighborhood R of P such that P is a cut point of the component C of $M \cdot \bar{R}$ which contains P. Hence $C - P = C_1 + C_2$, where C_1 and C_2 are mutually separated. By supposition there exists in M a simple closed curve J which contains P and lies wholly in R. Then clearly J must belong to C, and $J - P$ must belong either to C_1 or to C_2. But this is impossible, for since P is a point of order 2 of M, P must be a point of order 1 of each of the continua $C_1 + P$ and $C_2 + P$.

---

[9]) The subset R of a closed set M is said to be an open subset of M provided $M - R$ is either vacuous or closed.

[10]) A set M is uniformly connected im kleinen provided that for each $\varepsilon > 0$, a $\delta_\varepsilon > 0$ exists such that every two points x and y of M whose distance apart is $< \delta_\varepsilon$ lie in a connected subset of M of diameter $< \varepsilon$.

[11]) Bull. Amer. Math. Soc. **34** (1928), pp. 349–360.

Corollary. *Every im kleinen cycle point of M which is also an im kleinen cut point of M is a ramification point of M.*

Theorem 5. *If K, N and H respectively denote the set of all the im kleinen cut points, im kleinen cycle points, and end points of a continuous curve M, then $K+H+N=M$, $K \cdot N$ is countable, and $K \cdot H = N \cdot H = 0$.*

Proof. Let P be any point of M which belongs to neither K nor H. I shall show that P belongs to N. Let ε be any positive number. Since P does not belong to K, there exists a neighborhood R of P of diameter $< \varepsilon$ such that P is not a cut point of the component C of $M \cdot \bar{R}$ which contains P. Since P does not belong to H it follows[12]) that P is an interior point of some arc APB which belongs to C. And since P is a non-cut point of C it follows with the aid of a theorem of R. L. Moore's[13]) that $C-P$ contains an arc t from A to B. Clearly the sum of the arcs APB and t contains a simple closed curve J containing P; and since J must be of diameter $< \varepsilon$, it follows that P belongs to N. Therefore $K+H+N=M$.

By a theorem of the author's[14]) all save possibly a countable number of the points of K are points of order two of M. And by Theorem 4, no point of K which is a point of order two of M can belong to N. Hence $K \cdot N$ is countable. Obviously $K \cdot H = N \cdot H = 0$. This completes the proof of Theorem 5.

Corollary. *If M is cyclicly connected, then $H=0$ and $M=K+N$.*

Theorem 6. *Every point of a cyclicly connected continuous curve M which is not an im kleinen cycle point of M is a point of finite order of M.*

Proof. Let P be a non-im kleinen cycle point of M. There exists an $\varepsilon > 0$ such that P belongs to no simple closed curve in M of diameter $< \varepsilon$. Let R be an open set containing P and of diameter $< \varepsilon/4$, and let Q denote the component of $M \cdot \bar{R}$ which contains P. Since M is cyclicly connected, it follows that each component of $Q-P$ must contain at least one point of $F(R)$; and hence the components of $Q-P$ are finite in number. Let them be denoted by $E_1, E_2, \ldots, E_n$. I shall show that P is a point of order n of M. For each $i \leqq n$, P must be an

[12]) See footnote [5]) and Theorem 3 above. Although C itself is not necessarily a continuous curve, it follows by Theorem 3 that C contains a continuous curve U such that $U \supset P$ and $(M-U)' \cdot P = 0$.

[13]) Concerning Continuous Curves in the Plane, Math. Zeitschr. **15** (1922), Theorem 1. Also see footnote [12]).

[14]) G. T. Whyburn, Concerning Collections of Cuttings of Connected Point Sets, Bull. Amer. Math. Soc. **35** (1929), pp. 87–104.

endpoint (and hence a point of order 1) of the continuum $E_i + P$. For by a theorem due to W. L. Ayres and the author[15]) there exists a continuous curve W_i of diameter $< \varepsilon$ containing E_i, lying in M, and such that P is not a limit point of $W_i - (E_i + P)$. Then since P is not a cut point of W_i and lies on no simple closed curve in W_i, then[16]) P is an endpoint of W_i. Hence[17]) P is a point of order 1 of W_i and since P is not a limit point of $W_i - (E_i + P)$, P is a point of order 1 of $E_i + P$. And since P is a point of order 1 of each of the continua

$$E_1 + P, E_2 + P, \ldots, E_n + P,$$

it follows that P is a point of order n of M.

Theorem 7. *In order that the point P of a continuous curve M should be a non-im kleinen cycle point of M it is necessary and sufficient that there should exist a number $\varepsilon > 0$ such that P is an endpoint of each component into which P ε-cuts M (i. e., if R is a domain of diameter $< \varepsilon$ containing P and N is the component of $\bar{R} \cdot M$ containing P, then P is an endpoint of each continuum obtained by adding it to each component of $(N - P)$). Furthermore, if P ε-cuts M into n such pieces $(n = 1, 2, 3, \ldots)$ for some one ε, then P is a point of order n of M.*

Theorem 8. *The set K of all the im kleinen cut points of any continuous curve M is an F_σ (i. e., the sum of a countable number of closed sets).*

Proof. Let K_1 denote the set of all those points of K which are condensation points of K. It follows by Theorem 3 that for each point P of K_1 and each integer $n > 0$, M contains a continuous curve U_{pn} of diameter $< 1/n$ containing P and such that P is a cut point of U_{pn} but is not a limit point of $M - U_{pn}$ and such that every cut point of U_{pn} belongs to K. By the Lindelöf Theorem, for each n, there exists a countable subset N_n of K_1 such that $K_1 \subset \sum_{P \subset N_n} I(U_{pn})$, where $I(U_{pn}) =$ set of all inner points of U_{pn}. By a theorem of Zarankiewicz[18]) for each n and each $P \subset N_n$, the set K_{pn} of cut points of the curve U_{pn} is an F_σ, and hence $K_{pn} = \sum_{j=1}^{\infty} F_{jpn}$, where F_{jpn} is closed for every j. And if Q denotes the point set $K - K_1 + \sum_{n=1}^{\infty} \sum_{P \subset N_n} \sum_{j=1}^{\infty} F_{jpn}$, then since $K - K_1$ is

[15]) *Loc. cit.* See footnote [11]).

[16]) G. T. Whyburn, Concerning Continua in the Plane, *loc. cit.*; W. L. Ayres, Concerning Continuous Curves and Correspondences, Ann. of Math. **28** (1927), p. 396. For this theorem in n dimensions see W. L. Ayres, Concerning Continuous Curves in a Space of n Dimensions, *loc. cit.* [5]).

[17]) See footnote [5]).

[18]) C. Zarankiewicz, *loc. cit.*, Theorem 17.

countable it follows that Q is an F_σ. It remains to show that Q is identical with K. Clearly $Q \subset K$, for $K_{pn} \subset K$ for each n and each $P \subset N_n$. Let P be any point of K_1. If P belongs to any set N_n then it must belong to Q, for Q contains every N_n. If P belongs to no N_n, then P must be a limit point of a sequence of points $[P_n]$, where for each n, $P_n \subset N_n$. There exists a number $d > 0$ such that P is a cut point of the component R of $\overline{M \cdot S(P, 4d)}$ which contains P. There exists an integer $n_1 > 0$ such that $1/n_1 < d$ and a point P_1 of N_{n_1} such that P belongs to $U_{p_1 n_1}$ and is an inner point of $U_{p_1 n_1}$ relative to M. Clearly $U_{p_1 n_1} \subset R$ and since P is a cut point of R and is an inner point of $U_{p_1 n_1}$ it follows[19]) readily that P is a cut point of $U_{p_1 n_1}$. Therefore P belongs to $K_{p_1 n_1}$, and hence belongs to Q. Hence Q is identical with K, and therefore K is an F_σ.

Corollary. *The set of all the non-im kleinen cut points of any continuous curve M is a G_δ (relative to M), i. e., the common part of a family of sets each open in M.*

The following additional facts concerning the Borel classification of certain types of points of a continuous curve either are already known or are easily deduced. Let K^*, H, and N^* denote the set of all cut points, endpoints and points belonging to some simple closed curve in a continuous curve M, then

(I) K^* is an F_σ (theorem of Zarankiewicz, *loc. cit.*).

(II) N^* is an F_σ (theorem of the author's, *cf.* my paper *Cyclicly Connected Continuous Curves, loc. cit.*, where it is shown that N is the sum of a countable number of continuous curves, *i. e.*, the maximal cyclic curves of M).

(III) H is a G_δ (theorem of Menger, *loc. cit.*).

(IV) $M - K^*$, and $M - N^*$ are G_δ's.

Examples are easily constructed to show that K^* is not necessarily a G_δ. It would be interesting to determine whether or not (a) H is necessarily an F_σ and (b) N^* is necessarily a G_δ.

§ 3.

Density of the Non-Im Kleinen Cut Points and Ramification Points.

Theorem 9. *Let L denote the set of all non-im kleinen cut (avoidable) points of a continuous curve M and let t be any arc in M. Then if L is dense on t it is uncountably everywhere dense on t.*

---

[19]) In this connection see also, R. L. Moore, *loc. cit.* ref. [9]) Lemma 2.

Proof. By Theorem 8 and corollary, L is a G_δ relative to M. Hence if S is any arc segment in t, $S\cdot F$ is a G_δ relative to S. And since $S\cdot L$ is dense in S it follows by Young's Theorem[20]) that $S\cdot L$ has the power of the continuum. Hence L is uncountably dense on t.

Theorem 10. *Let T be any arc of a cyclicly connected continuous curve M and let L denote the set of all non-im kleinen cut points of M. Then if the set W of ramification points of M is dense on T, $L\cdot T$ has at least the power c of the continuum.*

Lemma 10a. *If T is any arc in a continuous curve M, then the set I of all points X of T such that for each $\varepsilon > 0$, M contains an ε-simple closed curve containing an arc segment in T which contains X, is a linear G_δ.*

Proof of Theorem 10. Suppose, on the contrary, that $L\cdot T$ has a cardinal number $< c$. Then by Theorem 9, L cannot be dense on T. Accordingly, there exists an interval t of T which contains no point of L. Then $t \subset K$. Let t_0 be any subarc of t of diameter $< 1/4$. Since by a theorem of the author's [see ref. [14])], $K\cdot W$ is countable, t_0 contains an interior point X_1 which is a point of order 2 of M. Since W is dense on t_0, X_1 is a limit point of $M - t_0$ but is not a limit point of any single component of $M - t_0$. Since M is a continuous curve, it is readily seen that there exists a component R_1 of $M - t_0$ of diameter $< 1/3$. As M is cyclicly connected, t_0 contains[21]) at least two limit points of R; and it is easily seen that $R_1 + t_0$ contains a simple closed curve J_1 of diameter < 1 which contains an interval t_1 of t_0 every point of which is interior to t_0 and which is of diameter $< 1/8$. Just as above it follows that M contains a simple closed curve J_2 of diameter $< 1/2$ which contains an interval t_2 of t_1 every point of which is interior to t_1 and which is of diameter $< 1/16$ and so on. Let this process be continued indefinitely. There exists a point X common to all of the intervals $t_0, t_1, t_2, \ldots$. Clearly X belongs to the set I (see Lemma 10a). And as t_0 is any interval of t, I must be dense on t. But by Lemma 10a, I is a G_δ. Hence by Young's Theorem above quoted, $I\cdot t$ is uncountable. Clearly $I\cdot t \subset N$ and since, by Theorem 5, $N\cdot K$ is countable, $I\cdot t$ must contain at least one point of L, contrary to the fact that $L\cdot t = 0$. Thus the supposition that Theorem 10 is false leads to a contradiction.

Corollary 1. *If the non-im kleinen cut points of a cyclicly connected continuous curve M are not uncountably dense on the arc t of M, then t contains an arc segment which is an open subset of M.*

---

[20]) *Cf.* W. H. Young, Leipz. Ber. **55** (1903), S. 287.

[21]) G. T. Whyburn, Cyclicly Connected Continuous Curves, *loc. cit.* [7]).

Corollary 2. *If the im kleinen cycle points of a cyclicly connected continuous curve M are dense on an arc t in M, then both the avoidable points and the im kleinen cycle points of M are uncountably everywhere dense on t.*

§ 4.

Curves Composed Almost Wholly of Im Kleinen Cut Points.

Theorem 11. *If every point of a continuous curve M is either an end point or an im kleinen cut point, then M is a Menger regular curve and both the ramification points and the im kleinen cycle points of M are countable.*

Proof. That M is a regular curve was proved by the author in another paper[22]). That the ramification points of M are countable follows from the fact that no endpoint is a ramification point and the author's theorem[22]) that only a countable number of the im kleinen cut points of M are ramification points; and that the im kleinen cycle points are countable follows from Theorem 5 and the fact that no end point is an im kleinen cycle point.

Theorem 12. *In order that every subcontinuum of a continuous curve M should contain an arc segment which is an open subset of M it is necessary and sufficient that if W denotes the set of all ramification points of M then $\overline{W}$ is totally disconnected.*

Proof. The condition is obviously necessary. It is also sufficient For let Q be any subcontinuum of M and let R be a component of $Q-\overline{W}\cdot Q$. Then[23]) R contains an arc AB, and since every point of AB is a point of order two of M, it follows that no point of $AB-(A+B)$ is a limit point of $M-[AB-(A+B)]$. Hence the segment AB is an open subset of M.

Theorem 13. *If every point of a cyclicly connected continuous curve M is an im kleinen cut point of M, and W denotes the set of all ramification points of M, then* (1) *$\overline{W}$ is totally disconnected, and* (2) *every component of $M-\overline{W}$ is an arc segment.*

Proof. Suppose, contrary to (1), that $\overline{W}$ contains a continuum C. Then since by Theorem of the author's mentioned above, M is a Menger regular curve, it follows[24]) that C is a regular curve, and hence C con-

---

[22]) *Cf.* G. T. Whyburn, Concerning Collections of Cuttings of Connected Point Sets, *loc. cit.*

[23]) R. L. Moore, *loc. cit.*, see footnote [13]).

[24]) K. Menger, *loc. cit.*, see footnote [8]).

tains an arc t. But then W is dense on t; and hence, by Theorem 10, t contains at least one point which is not an im kleinen cut point of M, contrary to hypothesis. Hence (1) is true.

Now let R be any component of $M-\overline{W}$. Since M is cyclicly connected, $\overline{W}$ contains[25]) at least two limit points of R. Hence it is readily seen that there exists an arc AB such that A and B belong to $\overline{W}$ and which lies except for the points A and B wholly in R. And since every point of the segment $AB-(A+B)$ is a point of order two of M, clearly R must be identical with this segment.

Essentially the same argument suffices to prove the following more general theorem.

Theorem 14. *If the avoidable points of a cyclicly connected continuous curve M are not uncountably dense on any subcontinuum of M, and if W denotes the set of all ramification points of M, then* (1) *$\overline{W}$ is totally disconnected, and* (2) *every component of $M-\overline{W}$ is an arc segment.*

Theorem 15. *If every subcontinuum of a bounded cyclicly connected continuous curve M contains an arc segment which is an open subset of M, and W denotes the set of all ramification points of M, then* (1) *$\overline{W}$ is totally disconnected,* (2) *M is a Menger regular curve,* (3) *each component of $M-\overline{W}$ is an arc segment,* (4) *if L is any closed totally disconnected subset of M containing $\overline{W}$, and G denotes the collection of components of $M-L$, then for each $\varepsilon>0$ every point of L can be ε-separated*[26]) *in M by a finite number of the segments of the collection G, and* (5) *for each $\varepsilon>0$, M is the sum of a finite number of mutually exlusive ε-continua plus a finite number of mutually exclusive ε-arc segments the two endpoints of each of which belong to different continua of the set just mentioned. Hence M is the sum of a finite number of ε-continua no two of which have more than one point in common*[27]).

Proof. Conclusions (1), (2), and (3) are obvious from the above theorems and discussion. To prove (4), let ε be any positive number and P any point of L. Since L is closed and totally disconnected, it follows

---

[25]) See footnote [21]).

[26]) For a given $\varepsilon>0$, the point P of M is said to be ε-separated in M by a set A provided $M-A=M_p+M_0$ where M_p and M_0 are mutually separated sets and M_p contains the point P and is of diameter $<\varepsilon$. *Cf.* P. Urysohn, Comptes Rendus **175** (1922), p. 481. Urysohn's definition differs from the one just given in that it requires that the set M_p+A be of diameter $<\varepsilon$.

[27]) K. Menger (Zur allgemeinen Kurventheorie, Fund. Math. **10**) has proposed the question as to whether or not every regular curve has the property mentioned in the last sentence of this theorem.

that there exists an open set R containing P and of diameter $< \varepsilon/2$ and such that $F(R)\cdot L = 0$. Then clearly $F(R)\cdot M \subset M - L = \sum_{g \in G} g$. Let G_p be the collection of all elements of G which contain at least one point of $F(R)$. Then G_p must be finite; for otherwise, since (see § 7 below) only a finite number of the segments of G are of diameter $>$ any positive number, it would follow that $F(R)$ contained a limit point of L, contrary to the fact that L is closed and $F(R)\cdot L = 0$. Hence G_p is finite and clearly G_p ε-separates P in M. And since G_p is finite, obviously it contains a subcollection which ε-separates P in M and is irreducible with respect to this property.

To prove (5), let ε be any number > 0. Since by (2) M is a regular curve, by a theorem of Menger's[28]) M is the sum of the elements of a finite collection Q of $\varepsilon/2$-continua each pair of which have at most a finite number of common points. Let E denote the set of all points which belong to at least two elements of the collection Q. The points of E are finite in number and can be ordered $P_1, P_2, P_3, \ldots, P_n$. Let L denote the set of points $\overline{W} + P_1 + P_2 + \ldots + P_n$. By (4), there exists a finite collection G_1 of the arc segments of G whose sum separates P_1 in M from $P_2 + P_3 + \ldots + P_n$ and which is irreducible with respect to this property. Each segment S of G_1 contains an arc segment U which lies, together with its end points, wholly in S; and if U_1 denotes the finite collection of segments U, it is easy to see that the sum of the segments of U_1 also separates P_1 in M from $P_2 + P_3 + \ldots + P_n$. Let V_1 denote the point set obtained by adding together all the point sets of the collection U_1. Now a similar argument shows that there exists a set V_2 which is the sum of the elements of a finite collection U_2 of arc-segments and which separates P_2 in M from $(P_1 + \overline{V}_1) + P_3 + P_4 + \ldots + P_n$; and indeed, for each i, $1 \leqq i \leqq n$, M contains a set V_i which is the sum of the elements of a finite collection U_i of arc segments selected as above and which separates P_i in M from $(P_1 + \overline{V}_1) + (P_2 + \overline{V}_2) + (P_{i-1} + \overline{V}_{i-1}) + P_{i+1} + P_{i+2} + \ldots + P_n$.

Now let U denote the collection of all the arc segments which belong to any collection U_i $(1 \leqq i \leqq n)$, and let V denote the point set $\sum_{i=1}^{n} V_i$. Let F denote the collection of all the components of $M - V$. Then each element of F is a continuum, and by a theorem proved by Kuratowski and Knaster[29]) and independently by the author[29]) it follows that F is

[28]) Grundzüge einer Theorie der Kurven, *loc. cit.*

[29]) *Cf.* Kuratowski and Knaster, Remark on a Theorem of R. L. Moore, Proc. Ntl. Acad. of Sci. **13** (1927); G. T. Whyburn, On the Separation of Connected Point Sets, Bull. Amer. Math. Soc. **33** (1927), p. 388 (abstract).

finite. And since no element of F can contain more than one point of the set E, each element of F must be of diameter $< \varepsilon$. And since each segment u in the collection U lies together with its end points in some segment of G and belongs to some collection U_i and hence is an element of an irreducible set of segments separating P_i in M from $(P_1 + \overline{V}_1) + \dots + (P_{i-1} + \overline{V}_{i-1}) + P_{i+1} + \dots + P_n$, it readily follows that u is of diameter $< \varepsilon$ and that not both end points of u belong to the same element of F. Hence (5) is true. And if for each u in U we let $X = \bar{u}$ and let D be the collection whose elements are the arcs X together with the elements of F, then clearly $M = \sum_D d$, and no two continua of D have more than one point in common. This completes the proof of Theorem 15.

Theorem 16. *If every point of the bounded cyclicly connected continuous curve M is an im kleinen cut point (or indeed if the non-im kleinen cut points of M are not dense on any subcontinuum of M), then M has properties* (1)–(5) *in Theorem* 15.

Theorem 17. *If every point of a bounded continuous curve M is an im kleinen cut point (or if the non-im kleinen cut points are not uncountably dense on any subcontinuum of M), then for each $\varepsilon > 0$, M is the sum of a finite number of ε-continua no two of which have more than one point in common.*

Proof. It follows by Theorem 16 that every maximal cyclic curve of M is for each ε, the sum of a finite number of ε-continua no two having more than one common point. Hence, by a theorem of the author's[30]) M itself has this same property.

Examples are easily constructed to show that (a) under the conditions of Theorem 15, neither W nor $M - K$ (the set of avoidable points) is necessarily countable, and (b) K (the im kleinen cut points) can be uncountably dense on every subcontinuum of a cyclicly connected continuous curve M and yet M not have property (1) in the statement of Theorem 15.

§ 5.

Node Curves.

Definition. A continuous curve M will be called a *node curve* provided that for each $\varepsilon > 0$, M is the sum of a finite number of ε-continua each having at most two points in common with the rest of M. It is obvious from this definition that every node curve is a bounded

[30]) G. T. Whyburn, Concerning Menger Regular Curves, Fund. Math. 12, Theorem 2.

Menger regular curve. However, as will be apparent below, the converse is not true. Hence a node curve is a special kind of a regular curve.

Theorem 18. *In order that a continuous curve M should be a node curve it is necessary and sufficient that for each $\varepsilon > 0$, M should contain a finite set of points Q such that each component of $M - Q$ is of diameter $< \varepsilon$ and has at most two limit points in Q.*

Theorem 19. *In order that the cyclicly connected continuous curve M should be a node curve it is necessary and sufficient that for each $\varepsilon > 0$, M is the sum of a finite number of ε-continua each having exactly two points in common with the rest of M. It is likewise necessary and sufficient that M contain a finite set Q such that each component of $M - Q$ is of diameter $< \varepsilon$ and has exactly two limit points in Q.*

Theorem 20. *In order that the bounded cyclicly connected continuous curve M should be a node curve without points of order w it is necessary and sufficient that no point of M of order > 2 (i. e., no ramification point) be an im kleinen cycle point.*

Proof. The condition is sufficient. It follows by Theorem 6 that every point of M is a point of finite order of M and hence M contains no point of order w. Let ε be any positive number which for convenience later we will suppose is $< 1/3$ the diameter of M. For each point X of M which is a point of order two of M there exist two points A_x and B_x of M and a connected open subset U_x of M containing P and of diameter $< \varepsilon$ and such that $\overline{U}_x \cdot (M - U_x) = A_x + B_x$. For each point Y of M which is not a point of order two of M, since Y is not an im kleinen cycle point of M, it follows as in the proof of Theorem 6 that a domain R exists containing Y, of diameter $< \varepsilon$, and such that Y cuts the component of $M \cdot \overline{R}$ which contains Y into n components (where n is the order of P) and is an end point (point of order one) of each of them. Thus it is readily seen that a connected open subset V_y of M exists which contains Y, is of diameter $< \varepsilon$, whose M-boundary contains just n points, and which is the sum of $Y + n$ open sets $U_{yx_1}, U_{yx_2}, \ldots, U_{yx_n}$ each having just Y and some other point A_{yx} as boundary points with respect to M. Let G_0 denote the collection of sets whose elements are the sets U_x and the sets V_x. Since G_0 covers M, then by the Borel Theorem G_0 contains a finite subcollection G which also covers M. By the above properties of the sets V_x, it follows readily that there exists a finite collection $U_1, U_2, \ldots, U_m$ of open subsets of M each of diameter $< \varepsilon$ and such that (1) for each i, $1 \leq i \leq m$, the M-boundary of U_i (*i. e.*, the boundary of U_i with respect to M) consists of just two points A_i and B_i, (2) $U_i + U_i \cdot U_j + U_j$ for each i and $j \leq m$, and (3) $M \subset (\overline{U}_1 + \overline{U}_2$

$+\overline{U}_2+\dots+\overline{U}_m)$. Let N denote the (finite) set of points $\sum_{i=1}^{m}(A_i+B_i)$. Then clearly each component of $M-N$ must be of diameter $<\varepsilon$, for it is a subset of some set $\overline{U}_i$. It remains to show that each component of $M-N$ has just two limit points in N. Suppose on the contrary that some component R of $M-N$ has as many as three limit points P_1, P_2, P_3 which belong to N.

Now clearly there exists an integer i, $1 \leqq i \leqq m$, such that $R \subset U_i$. Now since M is cyclicly connected it follows[31]) that every cut point of $\overline{U}_i$ must separate A_i and B_i in $\overline{U}_i$. And since $\overline{R}-R$ contains at least three points it readily follows that $\overline{R}-R$ contains at least one point P distinct from A_i and B_i which does not separate A_i and B_i in $\overline{U}_i$. Hence P is not a cut point of $\overline{U}_i$. Now since P belongs to N, there exists an integer j such that $A_j = P$. Now A_j belongs to U_i; and B_j cannot belong to U_i. For suppose $B_j \subset U_i$. Now $\overline{U}_i+\overline{U}_j \neq M$, for the diameter of M is $\geqq 3\varepsilon$. Let C be a component of $M-(\overline{U}_i+\overline{U}_j)$. Since M is cyclicly connected, $\overline{U}_i+\overline{U}_j$ must contain at least two limit points X_1 and X_2 of C. Obviously $X_1+X_2 \subset A_i+B_i+A_j+B_j$, and since $A_j+B_j \subset U_i$, then both A_i and B_i must be limit points of C. But since U_j contains a point of U_i but is not a subset of U_i, by (2), then U_j must contain at least one of the points A_i and B_i; and clearly this is impossible, since each of these points is a limit point of C. Therefore B_j does not belong to U_i. But now A_j is not a cut point of U_i. Hence U_i-A_j is connected and contains at least one point of U_j but contains neither A_j nor B_j; and therefore U_i is a subset of U_j, contrary to (2). Thus the supposition that some component of $M-N$ has more than two limit points in N leads to a contradiction; and hence, by Theorem 18, M is a node curve.

The condition is also necessary. For let M be any node curve containing no point of order ω, and let P be any point of M of order >2 of M. Let n be the order of P. Since M is a regular curve, by a theorem of Menger's[32]) there exists a set of n arcs $A_1P, A_2P, \dots, A_nP$ belonging to M and each having P as one end point but no two having in common any point except P. Let d be a number less than each of the numbers $\delta(A_1, P), \delta(A_2, P), \dots, \delta(A_n, P)$, where $\delta(A_i, P) =$ distance from A_i to P, and let R be a domain containing P and of diameter $< d/4$. Since M is a node curve, it is the sum of a finite collection G of continua each

---

[31]) With the aid of the following easily established lemma: *If the connected open subset R of a cyclicly connected continuous curve M has just two boundary points A and B with respect to M, then every cut point of the curve $\overline{R}$ separates A and B in $\overline{R}$ and if Q denotes the set of all such points, then $Q+A+B$ is closed.*

[32]) K. Menger, *loc. cit.*, see ref. [27]).

of diameter $< d/4$ each having at most two points in common with the rest of M. Since $n > 2$ it is readily seen that P is not an interior point of any one of the continua of G relative to M, and indeed, that there exist exactly n of the continua of G each of which contains P and some segment of one of the arcs $A_i P$ having P as one of its end points. Let $E_1, E_2, \ldots, E_n$ denote these continua. Then since P is a point of order n of M, it follows that P is a point of order 1 of each of the continua $E_1, E_2, \ldots, E_n$. Therefore, since P is not a limit point of $M - (E_1 + E_2 + \ldots + E_n)$, it follows that P is not an im kleinen cycle point of M. This completes the proof.

In proving the necessity of the condition in Theorem 20 no use was made of the fact that the curve M is cyclicly connected. Hence we have the following theorem.

Theorem 21. *No node curve M contains a point of finite order > 2 which is an im kleinen cycle point.*

The above proof for the sufficiency of the condition in Theorem 20 suffices to establish the following theorem.

Theorem 22. *Every bounded node curve im kleinen is a node curve.*

Theorem 23. *Every subcontinuum of a node curve is itself a node curve.*

Theorem 24. *If P is a point of order n (n finite and > 2) of a node curve M, then there exists a positive number ε_p such that if $\varepsilon < \varepsilon_p$, and M is decomposed into a finite collection G of ε-continua each having at most two points in common with the rest of M, then P is common to exactly n of the continua of G, P is an end point of each of these n continua and is not a limit point of M minus their sum.*

Proof. By Menger's theorem there exist n subarcs $A_1 P, A_2 P, \ldots, A_n P$ of M from A_1 to P, A_2 to $P, \ldots, A_n$ to P, respectively, each two having just the point P in common. There exists a hypersphere S with center P which neither contains nor encloses any of the points $A_1, A_2, \ldots, A_n$. Let ε_p denote 1/2 the radius of S. Let ε be any positive number $< \varepsilon_p$ and let M be decomposed as above into a set G of ε-continua. Let N denote the (finite) subset of M each point of which is common to at least two of the continua of G. On each of the arcs $A_i P$ ($i = 1, 2, \ldots, n$), in the order from P to A_i, let X_i denote the first point after P, which belongs to N. Now since $n > 2$, it is clear that P must belong to N. Each of the arc segments PX_i must lie wholly within one of the continua of G, and since each of these arc segments has at least one limit point X_i other than P which belongs to N, it is clear that no two of

these segments can lie in the same continuum of G. Hence P is common to at least n continua of G. And since P is of order n, clearly P belongs to no more than n such continua, is not a limit point of M minus the sum of these n continua, and is an end point of each.

A similar proof shows the following theorem.

Theorem 25. *If P is a point of order w of the node curve M, then for each integer $n > 0$ a number $\varepsilon_{np} > 0$ exists such that if $\varepsilon < \varepsilon_{np}$ and M is decomposed into a collection G of ε-continua as described above, then P is common to at least n of the continua of G.*

The two preceding theorems give the following theorem:

Theorem 26. *Every point of a node curve M of order > 2 of M is an im kleinen cut point of M.*

Theorem 27. *The ramification points of any node curve are countable.*

Theorem 27 follows immediately from Theorem 26 and the author's theorem that the im kleinen cut points of any continuous curve M of order > 2 are countable. It also follows from Theorems 24 and 25. For if for each integer $n > 0$, we decompose M into a finite collection G_n of $1/n$-continua as above, and N_n is the set of all points common to two of the continua of G_n, then by Theorems 24 and 25, the set W of points of M of order > 2 is a subset of $\sum_{n=1}^{\infty} N_n$. And since for each n, N_n is finite, $\sum_{n=1}^{\infty} N_n$ is countable, and hence W is countable.

Theorem 28. *The im kleinen cut points of every node curve M are everywhere dense in M. Indeed, every point of $\sum_{n=1}^{\infty} N_n$ above is an im kleinen cut point.*

Every maximal cyclic curve of a bounded continuous curve M can be a node curve and yet M itself not be a node curve, as seen by the following example. Let I denote the interval $(1, 2)$ of the X-axis, let C be the circle $X^2 + Y^2 = 1$, let R be the interior of C, and for each n $(n = 1, 2, 3, \ldots)$ let C_n denote the circle $(X-1)^2 + Y^2 = 1/n^2$. Then if M denotes the continuous curve $I + C + R \cdot \sum_{n=1}^{\infty} C_n$, the set of points $C + R \cdot \sum_{n=1}^{\infty} C_n$ is the only maximal cyclic curve of M; and although, as is easily seen with the aid of Theorem 20, this set of points is a node curve, nevertheless (*cf.* Theorem 21) M itself is not a node curve. However, we may characterize a node curve in terms of its maximal cyclic curves as follows.

Theorem 29. *In order that a bounded continuous curve M should be a node curve it is necessary and sufficient that for each maximal cyclic curve C of M (1) C is a node curve, and (2) every point of C which is a limit point of some component of $M - C$ is an im kleinen cut point of C.*

Proof. The conditions are necessary. Condition (1) is necessary because by Theorem 23, every subcontinuum of M is a node curve. To show that condition (2) is necessary, let C be any maximal cyclic curve of M and P any point of C which is a limit point of some component Q of $M - C$. There exists an arc T which has P as one end point and is a subset of $Q + P$. Suppose, contrary to what we purpose to show, that P is not an im kleinen cut point of C. Then since by condition (1), C is a node curve, it follows by Theorem 26 that P is a point of order two of C. Then P is a point of order 3 of the curve $C + T$, and since, by Theorem 23, $C + T$ is a node curve, then by Theorem 21, P is not an im kleinen cycle point of $C + T$. Hence P is not an im kleinen cycle point of C; but then by Theorem 4, corollary, P must be an im kleinen cut point of C, contrary to supposition. Thus the supposition that condition (2) is not necessary leads to a contradiction.

The conditions are also sufficient. For let M be any bounded continuous curve satisfying the conditions, let ε be any positive number, and let G denote the (finite) collection of all those maximal cyclic curves of M which are of diameter $> \varepsilon/4$. In my paper *Concerning Menger Regular Curves*[33]) it was shown that M contains a continuous curve Q containing all the curves of G and such that (1) Q is the sum of the curves of G plus a finite number $t_1, t_2, t_3, \ldots, t_m$ of simple cyclic chains[34]) of cyclic elements of M such that for each i, $(1 \leq i \leq m)$, t_i has at most one point in common with t_{i-1}, and (2) each component of $M - Q$ is of diameter $< \varepsilon/4$ and has just one limit point in Q. It is readily seen that there exists a finite number of cyclic elements $C_1, C_2, C_3, \ldots, C_n$ of M including all the curves of the collection G, and such that if U_i is any component of $Q - \sum C_i$, then $\overline{U}_i$ is a simple cyclic chain determined by an arc $A_i B_i$ in M and having at most the points A_i and B_i in common with $\sum C_i$. Clearly the components of $Q - \sum C_i$ are finite in number. Denote them by $U_1, U_2, U_3, \ldots, U_k$. Now for each i, $(1 \leq i \leq n)$, let E_i denote the (finite) set of points in C_i each of which is a limit point of

[33]) Fund. Math. 12, see proof of Theorem 2. For a statement of practically the same theorem see a forthcoming paper of W. L. Ayres entitled 'Concerning Arc Curves and Basic Subsets of a Continuous Curve. Second Paper'.

[34]) *Cf.* my paper Concerning the Structure of a Continuous Curve, *loc. cit.* [9]).

some component of $Q - C_i$, and let E denote the (finite) set of points $\sum E_i$. Now since E is finite and since, by hypothesis, each point of E is an im kleinen cut point of each curve of the set $C_1, C_2, C_3, \ldots, C_n$ which contains it, it follows that there exists a number $\delta > 0$ such that if P is any point of E, C_i is any curve of the collection $C_1, C_2, \ldots, C_n$, and R is any domain containing P and of diameter $< \delta$, then P is a cut point of the component of $C_i \cdot \bar{R}$ which contains P. Let d be a positive number $< \varepsilon/4$ and $< \delta/2$. Now since for each i, $(1 \leqq i \leqq n)$, C_i is a node curve or a point, it follows by Theorem 19 that for each i, a finite subset Q_i of C_i exists such that each component of $C_i - Q_i$ is of diameter $< d$ and has just two limit points in C_i. Now for each i, $(1 \leqq i \leqq k)$, as shown in my paper *Concerning Menger Regular Curves* (*loc. cit.*) there exists on the arc $A_i B_i$ a finite set of points V_i each of which separates A_i and B_i in M and such that each component of $\bar{U}_i - V_i$ is of diameter $< d$ and has at most two limit points in V_i. Let Z denote the (finite) set of points of the curve Q such that each point P of Q belongs to at least two of the chains $\bar{U}_1, \bar{U}_2, \ldots, \bar{U}_k$. Finally, let F denote the set of points $Z + E + \sum_{i=1}^{n} Q_i + \sum_{i=1}^{k} V_i$. Then F is finite, and it is not difficult to show that no component of $M - F$ has more than two limit points in F. Therefore, by Theorem 18, M is a node curve.

Corollary 1. *Every bounded acyclic continuous curve is a node curve.*

Corollary 2. *Every bounded baum im kleinen curve is a node curve.*

Corollary 3. *If every maximal cyclic curve of a bounded continuous curve M is a baum im kleinen (or, what is equivalent, contains only a finite number of simple closed curves) then M is a node curve.*

Theorem 30. *If no point of a bounded continuous curve M is an im kleinen cycle point, then M is a node curve.*

Theorem 30 is a corollary to Theorems 29 and 20.

It is easily shown by examples: (1) that the ramification points in a node curve M may be dense on some arc in M, and (2) that the ramification points may be dense on no subcontinuum of a continuous curve M and yet M not be a node curve.

Let M be a node curve, ε any positive number, and K a finite subset of M such that (see Theorem 18) each component of $M - K$ is of diameter $< \varepsilon$ and has at most two limit points in K. Let G denote the collection of sets obtained by adding to each component of $M - K$ its limit points in K. We shall call the elements of G the *links* of the

curve M. Those links of M which join on to the rest of M at two points will be called *ordinary links* of M, and those joining on at only one point will be called the *end-links* of M. By a *simple chain* X of links of M joining two links A and B of M is meant the sum of the elements of a finite collection E of links such that (1) X is a continuum (2) X contains both A and B (3) A and B are the end-links of the node curve X and are its only end-links, (4) X is irreducible with respect to the property of being the sum of the elements of a finite collection of links of M and having properties (1) — (3). The following propositions are easily deduced:

a) Every two links of a connected set H of links of M can be joined by a simple chain of links lying in H;

b) The ordinary links of M are finite in number and their sum is a continuum;

c) If X is any simple chain of links of M between two links A and B, X_A and X_B are points of $M-K$ belonging to A and B respectively, then every point of $[X-(A+B)]\cdot K$ separates X_A and X_B in X. And every link of X except A and B separates X_A and X_B in X;

d) If M is cyclicly connected, all its links are ordinary links.

§ 6.

Im Kleinen Cut Points and Irreducible Cuttings.

If A and B are points of a continuum M and K is a subset of M such that $M-K$ is the sum of two mutually separated sets M_a and M_b containing A and B respectively, then K is said to be a cutting of M between A and B; if no proper subset of K is a cutting of M between A and B, K is called an *irreducible cutting*[35]) *of M between A and B*.

Theorem 31. *In order that the point P of a continuous curve M should be an im kleinen cut point of M it is necessary and sufficient that P be an isolated point of some irreducible cutting of M between some two points A and B of M.*

Proof. The condition is sufficient. For suppose P is an isolated point of an irreducible cutting K of M between the points A and B of M. Let R be a domain containing P and such that $\bar{R}$ contains no point of $A+B+K-P$, and let N be the component of $M\cdot\bar{R}$ containing P. Then since[36]) P is a limit point of both the components R_a and R_b of

---

[35]) *Cf.* G. T. Whyburn, Concerning Irreducible Cuttings of Continua, Fund. Math. **13**, pp. 42—57.

[36]) G. T. Whyburn, *loc. cit.*, Theorem 7.

$M-K$ containing A and B respectively, then N contains points of both R_a and R_b; and since $\bar{R}\cdot K=P$, it is readily seen that $N-P=N\cdot R_a+N\cdot R_b$, and $N\cdot R_a$ and $N\cdot R_b$ are mutually separated. Hence P is a cut point of N and is then an im kleinen cut point of M.

The condition is also necessary. For if P is an im kleinen cut point of M, a domain R exists such that P is a cut point of the component N of $M\cdot\bar{R}$ which contains P. Let A and B be points of N lying in R and belonging to different components of $N-P$ and such that there exists an arc AB in $M\cdot R$. Now clearly $P+F(R)\cdot M$ cuts M between A and B. Then[37]) $P+F(R)\cdot M$ contains an irreducible cutting K of M between A and B. And since there exists an arc AB in $M\cdot R$, and $K\cdot R\subset P$, it is clear that P must belong to K and must be an isolated point of K.

Theorem 32. *If K denotes the set of all the im kleinen cut points of any Menger regular curve M, then $M-K$ is totally disconnected.*

Proof. Suppose, on the contrary, that $M-K$ contains a connected set H containing two distinct points A and B. Since M is a regular curve, there exists a finite cutting Q of M, between A and B. By a theorem of the author's[38]) Q contains an irreducible cutting Q_0 of M between A and B. Since H is a connected subset of M containing both A and B, obviously it must contain at least one point P of Q_0. And since Q_0 is finite, P is an isolated point of Q_0. But then by Theorem 31, P must belong to K, contrary to the fact that $P\subset H$ and $H\cdot K=0$.

Corollary. *Under the hypothesis of Theorem 32, K is dense on every connected subset of M.*

Theorem 32 does not remain true if the hypothesis that "M is a Menger regular curve" is replaced by the weaker one that "every subcontinuum of M is a continuous curve". This fact is demonstrated in an example due to H. M. Gehman[39]).

§ 7.

Continua All of Whose Subcontinua Are Continuous Curves and Menger Regular Curves, in n Dimensions.

In this section I shall first give a simple example in 3-space of a continuous curve every subcontinuum of which is a continuous curve and which has some rather interesting properties. Referring to a system of

---

[37]) G. T. Whyburn, *loc. cit.*, Theorem 8.

[38]) *Loc. cit.*, Theorem 8.

[39]) Concerning the Subsets of a Plane Continuous Curve, Annals of Math. 27 (1925), pp. 29–46.

cylindrical coordinate axes $\mathsf{P}, \Theta, \mathsf{Z}$ in 3-space, let AB be the interval $(0, 1)$ of the Z-axis, let n take on in ascending order the set of values included in the set of all positive prime integers. For each n, let us subdivide AB into a set I_n of n equal subintervals by inserting a set K_n of $n-1$ points of subdivision; and in the plane $\Theta = \pi/n$, let us construct on each interval of the set I_n a semicircle having this interval as its diameter, and let C_n be the sum of all these semicircles. Let ψ denote the continuum

$$AB + \sum C_n.$$

Properties of the curve ψ:

(α) *Every subcontinuum of ψ is a continuous curve.*

This property is obvious from the construction; or it can easily be proved with the aid of a theorem of H. M. Gehman's[40]).

(β) *ψ is not a Menger regular curve.*

This property readily follows from property (γ) below.

(γ) *ψ contains infinitely many arcs AX_iB $(i = 1, 2, \ldots)$ from A to B no two having any common point except A and B.* (γ') *There exists an $\varepsilon > 0$ such that ψ contains infinitely many mutually exclusive continua of diameter $> \varepsilon$.*

To show property (γ) it is only necessary to set $AX_iB = C_i$ for each i.

(δ) *ψ contains an arcwise connected set N which is not arcwise connected im kleinen and which has a boundary point P which is not accessible from it.*

To prove (δ), we merely set $N = \sum C_n$ and let P be the point $(0, 0, 1/2\sqrt{2})$. Clearly N is not arcwise connected im kleinen at any one of its points belonging to $AB - (A + B)$, and obviously P is not accessible from N.

ε) *ψ contains a connected subset which is not arcwise connected.*

The set $N + P$ under (δ) is not arcwise connected, and hence ψ has property (ε).

(ζ) *ψ contains an arcwise connected and connected im kleinen set which is not arcwise connected im kleinen.*

The set N defined under (δ) satisfies all requirements on the set in (ζ). That this set, contrary to a statement made in the first abstract of this paper, is connected im kleinen was kindly pointed out to me by Professor R. L. Wilder. It would be interesting to determine whether a set N could

---

[40]) Some Conditions Under Which a Continuum is a Continuous Curve, Ann. of Math. 27 (1926), pp. 381–384, see Theorem 2.

lie in the plane and have property (ζ), or slightly different, whether a set N could lie in the plane and be strongly connected and connected im kleinen and yet not be strongly connected im kleinen.

Gehman[41]) has given an example of a plane continuum having both properties (α) and (β) and has shown that (α') and the absence of (γ') are equivalent for bounded continua in the plane. Zarankiewicz[42]) characterizes a continuum (in n-space) having property (α) as one containing no "continuum of convergence", and in footnote states that his condition is equivalent to the absence of (γ'), (*i. e.* to Gehman's condition). The curve ψ above, of course, shows although this is true in the plane that this is not the case in n-space for $n > 2$. It has been shown by the author[43]) that no plane continuum can have both properties (α) and (δ). Knaster and Kuratowski[44]) have given an example of a plane regular curve having property (ε). Their example is somewhat more complicated.

Theorem 33. *If M is a bounded Menger regular curve and ε is any positive number, then M does not contain more than a finite number of mutually exclusive continua each of diameter $> \varepsilon$.*

Proof. Suppose, on the contrary, that M contains an infinite sequence of continua $M_1, M_2, M_3, \ldots$ all of diameter $> \varepsilon$. Then since M is bounded, there exist two points A and B of M belonging to the limiting set of the sequence $M_1, M_2, \ldots$. But since M is a regular curve, there exists a finite subset K of M which separates A and B in M. Clearly this is impossible, since only a finite number of the continua $M_1, M_2, \ldots$ can contain points of K. Thus the supposition that Theorem 33 is false leads to a contradiction.

Theorem 34. *If M is any continuum having the property that for each $\varepsilon > 0$, M does not contain infinitely many mutually exclusive continua each of diameter $> \varepsilon$, then every subcontinuum of M is a continuous curve, and if H is any arcwise connected subset of M, then* (1) *H is arcwise connected im kleinen and* (2) *every boundary point P of H is regularly accessible*[45]) *from H.*

---

[41]) See reference [39]).

[42]) Sur les Points de Division dans les Ensembles Connexes, *loc. cit.*; K is a continuum of convergence of a continuum M provided that $M - K$ contains a sequence of continua whose sequential limiting set is K.

[43]) Concerning Certain Types of Continuous Curves, Proc. Ntl. Acad. of Sci. **12** (1926), pp. 761–767.

[44]) Knaster and Kuratowski, Bull. Amer. Math. Soc. **33** (1927), p. 106.

[45]) That is, for each $\varepsilon > 0$, a $\delta_\varepsilon > 0$ exists such that every point of H whose distance from P is $< \delta_\varepsilon$ can be joined to P by an arc in $H + P$ of diameter $< \varepsilon$. See my paper Concerning the Open Subsets of a Plane Continuous Curve, Proc. Ntl. Acad. of Sci. **13** (1927), pp. 650–656.

That every subcontinuum of M is a continuous curve follows by the proof given by Gehman (*loc. cit.*) for the case of the plane, which extends to n dimensions without difficulty. That M has property (1) follows by the proof given in my paper "Concerning Certain Types of Continuous Curves"[46]) for the theorem that every arcwise connected subset of a plane continuum every subcontinuum of which is a continuous curve is arcwise connected im kleinen, in the proof of which the only plane property used is that for each $\varepsilon > 0$ the continuum contains not more than a finite number of mutually exclusive continua each of diameter $< \varepsilon$. That M has property (2) follows in a similar way by a proof given in my paper "Concerning the Complementary Domains of Continua"[47]) for the theorem that every boundary point of an arcwise connected subset of any plane continuous curve every subcontinuum of which is a continuous curve is regularly accessible from that set.

Theorem 35. *If H is any arcwise connected subset of a Menger regular curve M, then* (1) *H is arcwise connected im kleinen,* (2) *every boundary point of H is regularly accessible from H.*

Theorem 35 is an immediate consequence of Theorems 33 and 34.

Theorem 36. *If K is any closed subset of a bounded continuum M every subcontinuum of which is a continuous curve, then for each $\varepsilon > 0$, $M - K$ contains at most a finite number of components each of diameter $> \varepsilon$.*

Theorem 36 follows at once with the aid of Zarankiewicz's Theorem (*loc. cit.*) that no continuum every subcontinuum of which is a continuous curve can contain a continuum of convergence.

Theorem 37. *If R is any connected open subset of a bounded continuous curve M every subcontinuum of which is a continuous curve, then R has property S*[48]) *and every boundary point of R is regularly accessible from R.*

Proof. Suppose theorem 37 is not true. Then either R does not have property S or R has a boundary point which is not regularly accessible from R. In either case[49]) it follows that there exists an $\varepsilon > 0$

---

[46]) *Loc. cit.*, see reference [43]).

[47]) Ann. of Math. **29** (1928), pp. 399–411.

[48]) That is, for each $\varepsilon > 0$, R is the sum of a finite number of connected sets each of diameter $< \varepsilon$. See R. L. Moore, Fund. Math. **3** (1922), p. 232.

[49]) For the former case, see my paper Concerning the Open Subsets of a Continuous Curve, *loc. cit.*, proof of Theorem 1, p. 651; and for the latter case see my paper Concerning Menger Regular Curves, Fund. Math. **12** (1928), Fundamental Accessibility Theorem.

and R contains an infinite sequence of points $P_1, P_2, \ldots$ having a sequential limit point P belonging to $M - R$ and such that no two of these points can be joined by any arc in R of diameter $< \varepsilon$. Let C be a circle with the center P and diameter $\varepsilon/2$. For each i, R contains an arc $P_i P_{i+1}$. Since each such arc must contain points without C, it readily follows that R contains a sequence of mutually exclusive arcs $T_1, T_2, T_3, \ldots$ each of diameter $> \varepsilon/8$ such that there exists a continuum T belonging to $M - R$ and which is the sequential limiting set of the sequence of continua $T_1, T_2, \ldots$. But then T is a continuum of convergence of M, contrary to a theorem of Zarankiewicz (*loc. cit.*). This contradiction proves Theorem 37.

(Eingegangen am 17. 11. 1928.)

Topologische Begründung des Kalküls der abzählenden Geometrie.

Von

Bartel L. van der Waerden in Groningen (Niederlande).

§ 1.

Einleitung.

Eines der Pariser Probleme Hilberts[1]) lautet: „Eine strenge Begründung des Schubertschen Abzählungskalküls".

In früheren Arbeiten[2]) habe ich gesucht darzutun, daß das Kernproblem der abzählenden Geometrie besteht in der Aufstellung einer brauchbaren Definition der „Multiplizitäten" oder der Vielfachheiten, mit denen die Lösungen eines algebraisch-geometrischen Problems gezählt werden müssen, damit das „Prinzip der Erhaltung der Anzahl" für diese Lösungen bei jeder Spezialisierung der Daten des Problems gelte. In der Arbeit W_1 habe ich gezeigt, daß man bei jedem Problem, dessen Gleichungen homogen in den Unbekannten und rational in einigen Parametern sind, die Lösungen für jede spezielle Parameterzahl in einer und nur einer Weise mit solchen Vielfachheiten versehen kann, daß Anzahl und algebraische Eigenschaften der Lösungen bei diesen Parameterspezialisierungen erhalten bleiben, und daß bei allgemeiner Parameterwahl die Multiplizitäten gleich 1 sind. Damit war eine implizite Definition der Multiplizitäten gegeben, aber noch kein brauchbares Mittel, diese in vorliegenden Fällen (außer den allereinfachsten) wirklich zu bestimmen. Eine besondere Schwierigkeit bei der Anwendung war noch, daß mit der Möglichkeit von „Lösungen mit der

---

1) D. Hilbert, Mathematische Probleme, Gött. Nachr. 1900, S. 253.

2) B. L. v. d. Waerden, Diss. Amsterdam 1926. Der Multiplizitätsbegriff der algebraischen Geometrie, Math. Annalen **97** (1927), S. 756 (*zitiert* W_1). Eine Verallgemeinerung des Bézoutschen Theorems, Math. Annalen **99** (1928), S. 497 (*zitiert* W_2). On Hilbert's function etc., Proc. Kon. Ak. Amsterdam **31** (1928), S. 749.

Multiplizität Null“, also von Lösungen, die bei spezieller Parameterwahl vorhanden sind, aber denen im allgemeinen Fall nichts entspricht, gerechnet werden mußte.

Die für den Schubertschen „Bedingungskalkül“ [3]) wichtigen Fälle sind die, wo es sich darum handelt, die Anzahl der gemeinsamen Elemente von zwei oder mehr algebraischen Varietäten [4]) in einer festen singularitätenfreien Mannigfaltigkeit (z. B. im komplexen projektiven Raum) zu bestimmen. Für den Fall einer Kurve V_1 und einer Hyperfläche V_{n-1} im projektiven R_n sind die Bestimmungen der Vielfachheiten durch explizite Formeln und die Berechnung der Gesamtzahl sowohl auf funktionentheoretischem als auf idealtheoretischem Wege gelungen [5]). Beide Bestimmungsweisen scheiterten jedoch gänzlich im allgemeineren Fall der Schnittpunkte einer V_r und V_{n-r} im projektiven R_n [6]). Die oben dargestellte implizite Multiplizitätsdefinition führte aber in diesen Fällen noch zum Ziel [7]). Dadurch nämlich, daß die Varietäten V_r und V_s mittels einer projektiven Transformation mit unbestimmten Koeffizienten in allgemeine Lage zueinander gebracht wurden, war das Schnittpunktsproblem abhängig gemacht von den Transformationsparametern, und die obige Definition der Multiplizität wurde anwendbar. Es gelang durch eine äußerst mühevolle Analyse, für spezielle Werte der Parameter („ausgeartete Transformation“) Anzahl und Multiplizitäten der Schnittpunkte algebraisch zu bestimmen, wobei sich das erwartete Ergebnis: Summe der Multiplizitäten = Produkt der Gradzahlen ergab, welches Ergebnis sich dann vermöge der „Erhaltung der Anzahl“ auf die allgemeine Lage sowie auf alle überhaupt möglichen Spezialisierungen übertrug.

Soweit sie reichte, hatte die algebraische Methode eine größere Allgemeinheit als jede analytische, da sie auf beliebige abstrakte Geometrien (die zu abstrakten Körpern gehören) anwendbar war. Aber bei der Übertragung der Methode auf Varietäten von Geraden u. dgl. stieß die Durchführung der Beweise auf immer wachsende Schwierigkeiten, und für solche Gebilde, die nicht wie der projektive Raum eine transitive Gruppe von Transformationen in sich gestatten, ist die Übertragung der obigen Multiplizitätsdefinition ganz ausgeschlossen.

---

[3]) H. Schubert, Kalkül der abzählenden Geometrie, Leipzig 1879.

[4]) Da das Wort *Mannigfaltigkeit* in dieser Arbeit, dem topologischen Sprachgebrauch entsprechend, für *singularitätenfreie* Räume reserviert bleibt, werde ich für die durch algebraische Gleichungen definierten Punktmengen das französische Wort „*Varietäten*“ gebrauchen.

[5]) Für Literatur siehe W_2, Einleitung.

[6]) W_2, Einleitung und § 11.

[7]) W_2.

Auch die analytischen Methoden [Zeuthen[8]), Halphen[9]) u. a.] erreichen nur bestimmte Fälle.

Aber die Topologie besitzt einen Multiplizitätsbegriff: den Begriff des Index eines Schnittpunktes von zwei Komplexen[10]), der schon von Lefschetz[11]) mit Erfolg auf die Theorie der algebraischen Flächen sowie auf Korrespondenzen auf algebraischen Kurven angewandt wurde. Soll dieser Indexbegriff, angewandt auf algebraische Varietäten im komplexen Gebiet, sich als Multiplizitätsbegriff für die abzählende Geometrie eignen, so muß er die folgenden drei Eigenschaften besitzen:

1. Die Summe der Indizes soll dem „Prinzip der Erhaltung der Anzahl" bei stetigen Änderungen der schneidenden Varietäten genügen.

2. Die Indizes sollen für einfache Schnittpunkte (d. h. wenn die Tangentialräume der schneidenden Varietäten nur einen Punkt gemein haben) alle den Wert 1 haben.

3. Sie sollen keine negativen Werte annehmen.

Die erste Eigenschaft des topologischen Indexbegriffs folgt fast unmittelbar aus seiner Definition durch simpliziale Approximationen. Die zweite Eigenschaft wurde allgemein für analytische Varietäten von Lefschetz[12]) bewiesen. Und über die dritte Eigenschaft hinaus läßt sich sogar zeigen, daß die Indizes bei analytischen (also insbesondere algebraischen) Varietäten immer *positiv* sind, wodurch also zugleich das unangenehme Vorkommnis der „Multiplizität Null" ausgeschlossen wird. Dieser Satz über analytische Varietäten im komplexen Gebiet ist das Hauptergebnis dieser Arbeit. Aus den Eigenschaften 1. bis 3. folgt dann unschwer die Übereinstimmung des topologischen mit dem algebraischen Multiplizitätsbegriff für den projektiven Raum.

Die Topologie leistet aber noch mehr als die Ermöglichung einer brauchbaren Multiplizitätsdefinition. Sie verschafft zugleich eine Fülle von Mitteln, die Indexsumme aller Schnittpunkte oder „Schnittpunktszahl", deren Bestimmung das Ziel aller abzählenden Methoden ist, in einfacher Weise zu bestimmen, indem sie zeigt, daß diese Indexsumme nur von den

---

[8]) H. G. Zeuthen, Abzählende Methoden (Leipzig 1914); Enzyklopädie III, 3.

[9]) Halphen, Comptes Rendus (4. Sept. 1876).

[10]) S. Lefschetz, Trans. Am. Math. Soc. 28 (1926), S. 1 (*zitiert* L_1).

[11]) S. Lefschetz, L'analysis situs et la géométrie algébrique, Paris 1924.

[12]) loc. cit. [11]), S. 19. Lefschetz schließt aus den Eigenschaften 1., 2. auf die (zur Positivität verschärfte) Eigenschaft 3. mit der Begründung, man könne durch eine kleine Verschiebung immer die unter 2. vorausgesetzte Situation herstellen. Die Begründung scheint in dieser Form ungenügend, denn bei der Verschiebung könnten ja Schnittpunkte in Wegfall kommen. Erst die genauere Betrachtung von § 5 wird lehren, daß dieses tatsächlich nicht vorkommen kann.

Homologieklassen der zum Schnitt gebrachten Varietäten abhängt, und indem sie für die Bestimmung der Homologieklassen den ganzen Apparat der „kombinatorischen Topologie“ zur Verfügung stellt. Zum Beispiel reduziert sich der algebraisch so mühevoll bewiesene Satz, daß im projektiven Raum die Schnittpunktsanzahl einer V_r und einer V_{n-r} gleich dem Produkt der Gradzahlen ist, von topologischem Gesichtspunkt auf die beiden leicht zu beweisenden Tatsachen, daß eine V_r (deren topologische Dimensionszahl im Komplexen gleich $2r$ ist) homolog dem g-fachen eines linearen L_r ist, wo g der Grad ist, und daß die Schnittpunktszahl zweier linearer Räume L_r und L_{n-r} gleich 1 ist.

Allgemein ergibt jede Homologierelation zwischen algebraischen Varietäten eine symbolische Gleichung im Schubertschen Sinn, und man darf diese Gleichungen unbeschränkt addieren und multiplizieren, wie es im Schubertschen Kalkül geschieht. Aus der Existenz einer endlichen Basis für die Homologien in jeder geschlossenen Mannigfaltigkeit ergibt sich weiter allgemein die Lösbarkeit der Schubertschen „Charakteristikenprobleme“.

Durch vollständige Angabe aller benutzten topologischen, analytischen und algebraischen Definitionen und Zusammenstellung der wichtigsten Sätze hoffe ich, die Schwierigkeit, die darin besteht, daß eine gewisse Vertrautheit sowohl mit topologischen als auch mit abzählenden Methoden beim Leser notwendig vorausgesetzt werden mußte, zu einem Minimum reduziert zu haben. Der sachverständige Leser muß dafür einige Ausführungen über bekannte Tatsachen mit in den Kauf nehmen. Die Kenntnis der zitierten Arbeiten W_1 und W_2 ist für das Verständnis dieser Arbeit nicht erforderlich. Einige topologische Hilfsbetrachtungen fast trivialer Natur sind, um den Gedankengang nicht zu unterbrechen, in zwei Anhängen vereinigt.

Anwendungen der hier zu entwickelnden Methoden auf konkrete abzählende Probleme hoffe ich später zu geben.

§ 2.

Komplexe im euklidischen Raum.

In diesem Paragraphen sollen, um allen Zweifel beim Gebrauch von Worten mit schwankender Bedeutung auszuschließen, die nötigen topologischen Grundbegriffe ganz kurz zusammengestellt werden. Für genauere Erörterungen sei auf die Lehrbücher verwiesen[13]).

---

[13]) Etwa: O. Veblen, The Cambridge Colloquium Lectures, Cambridge (Mass.) 1922. Oder Hadamards Note zu J. Tannery, Introduction à la théorie des fonctions II, zu ergänzen durch J. W. Alexanders Proof of the Invariance of certain Numbers, Proc. Am. Math. Soc. **16** (1915), p. 148.

Ein *geradliniger k-dimensionaler Komplex* ist aus endlichvielen k-dimensionalen euklidischen Simplizes aufgebaut. Ein stetiges Bild eines geradlinigen k-dimensionalen Komplexes im euklidischen R^n (mit ganz beliebigen Singularitäten) heißt *Komplex* schlechthin und wird mit K^k bezeichnet. Der leere Komplex heißt *Null*. Ist von Rand, von Orientierung, Unterteilung oder Addition von Komplexen die Rede, so ist damit immer gemeint, daß man die genannten Operationen zunächst an den Urbildern vornimmt und dann auf das Bild überträgt. *Orientierung eines Simplex* (im Urbild) geschieht dadurch, daß man einer bestimmten Reihenfolge der Ecken und allen geraden Permutationen davon ein Vorzeichen $\varepsilon = \pm 1$, allen ungeraden Permutationen das entgegengesetzte Vorzeichen $-\varepsilon$ zuordnet. *Orientierung eines Komplexes K^k* geschieht durch (beliebige) Orientierung aller seiner k-dimensionalen Simplizes. Bei der *Addition von Komplexen* wird jedes Simplex, welches in der Summe zweimal mit entgegengesetzter Orientierung vorkommt, diese beiden Male weggelassen. Mit jeder Orientierung eines Simplex ist nach einer bestimmten Vorschrift eine bestimmte Orientierung seines Randes verknüpft. Wie diese Vorschrift lautet, ist gleichgültig; nur muß sie so eingerichtet werden, daß der orientierte Rand des orientierten Randes gleich Null wird. Unter dem Rand $R(K^k)$ eines orientierten Komplexes K^k ist zu verstehen die Summe der orientierten $(k-1)$-dimensionalen Ränder der Simplizes von K^k. Zwei Komplexe K_1^r, K_2^r heißen *homolog zueinander im Gebiet U* (Gebiet heißt in dieser Arbeit eine beliebige offene Menge des R^n), wenn es einen K^{r+1} in U gibt derart, daß $K_1^r = K_2 + R(K^{r+1})$ ist. Zeichen für Homologie: $K_1^r \sim K_2^r$.

Komplexe mit Rand Null heißen *Zyklen*.

Eine (geradlinige) *δ-Approximation* eines Komplexes K^s entsteht dadurch, daß man, von einer hinreichend feinen Simplexeinteilung des Urbildes von K^s ausgehend, die den Bildpunkt K^s definierenden Abbildungsfunktionen durch solche ersetzt, die in jedem Simplex des Urbildes (Rand eingeschlossen) linear sind, derart, daß jeder Bildpunkt nach dieser Approximation zu seiner ursprünglichen Lage eine Entfernung $< \delta$ hat, und daß solche Simplizes, die eine Seite gemein haben, auch nach der Approximation die entsprechende Seite gemein haben.

Verbindet man jeden Punkt von K^s geradlinig mit dem entsprechenden Punkt des approximierenden Komplexes K_1^s, so erhält man einen „Verbindungskomplex" K_v^{s+1} in der δ-Umgebung von K^s und ebenso einen Randverbindungskomplex K_v^s in der δ-Umgebung des Randes $R(K^s)$ mit folgenden Eigenschaften:

$$R(K_v^{s+1}) = K^s - K_1^s - K_v^s,$$

$$R(K_v^s) = R(K^s) - R(K_1^s),$$

$$K_v^s = 0, \text{ wenn } R(K^s) \text{ Null oder geradlinig ist.}$$

Eine bestimmte Reihenfolge $x_1, \ldots, x_n$ der Cartesischen Koordinaten des R^n induziert eine bestimmte Orientierung aller n-dimensionalen geradlinigen Simplizes dieses Raumes (oder, wie wir auch sagen wollen, eine *Orientierung des Raumes* R^n), nämlich diejenige, wobei einer bestimmten Reihenfolge der Ecken $x^0, \ldots, x^n$ eines solchen Simplex als Vorzeichen zugeordnet wird das Vorzeichen der Determinante

$$|1\, x_1^\nu, \ldots, x_n^\nu| \quad \text{oder} \quad |x_1^\nu - x_1^0, \ldots, x_n^\nu - x_n^0|.$$

§ 3.

Schnitte von Komplexen im euklidischen R^n.

Der Begriff des Schnittkomplexes zweier Komplexe ist in der hier gebrauchten Form vor allem von Lefschetz (L_1) entwickelt worden. Eine Übersicht über die benötigten Begriffsbildungen folgt in diesem Paragraphen.

Sind zwei (geradlinige) Simplizes X^r und X^s im R^n gegeben, so kann man durch beliebig kleine Verrückungen der Ecken eine *„allgemeine Lage"* der Simplizes zueinander erzwingen, welche darin bestehen soll, daß die Räume R^r, R^s, in denen sie liegen, einen linearen Raum R^k von genau $k = r + s - n$ Dimensionen gemein haben (bzw. zueinander fremd sind für $r + s - n < 0$), und daß in diesem Schnittraum die beiden Simplizes entweder fremd sind oder innere Punkte gemein haben. Im letzten Fall haben sie ein konvexes Polyeder P^k von der Dimension k gemein, das wir irgendwie in Simplizes eingeteilt denken.

Orientieren wir die Räume R^n, R^r, R^s (oder den Raum R^n und die Simplizes X^r, X^s), so ist dadurch eine Orientierung eines jeden Simplex von P^k mitbestimmt nach folgender *Vorschrift*: Man ergänze die Ecken $x^0, \ldots, x^k$ eines Simplex von P^k durch Punkte $y^1, \ldots, y^{r-k}$ zu einem Eckpunktssystem eines Simplex von R^r; dessen Vorzeichen in der Orientierung von R^r sei ε_1. Ebenso ergänzen wir $x^0, \ldots, x^k$ mit $z^1, \ldots, z^{s-k}$ in R^s und bestimmen das Vorzeichen ε_2. Dann bilden die $n+1$ Punkte $x^0, \ldots, x^k, y^1, \ldots, y^{r-k}, z^1, \ldots, z^{s-k}$ die Ecken eines Simplex in R^n. Das zugehörige Vorzeichen in der Orientierung von R^n sei ε_3. Nunmehr ordnen wir der Eckenfolge der $x^0, \ldots, x^k$ das Vorzeichen

$$\varepsilon = \varepsilon_1 \varepsilon_2 \varepsilon_3$$

zu. Dadurch sind alle Simplizes von P^k und mithin auch P^k selbst orientiert.

Man nennt das so orientierte P^k den *orientierten Schnittkomplex* der Simplizes X^r und X^s und schreibt

$$P^k = X^r \cdot X^s.$$

Ist P^k leer, so schreibt man

$$X^r \cdot X^s = 0.$$

Für zwei beliebige geradlinige Komplexe K^r, K^s definieren wir nun, falls alle Simplizes von K^r (sowie ihre Seiten) zu denen von K^s die oben präzisierte „allgemeine Lage" haben, den *orientierten Schnittkomplex* $K^r \cdot K^s$ als die Summe der orientierten Schnitte der Simplizes von K^r mit denen von K^s.

Die wichtigsten Rechnungsregeln für orientierte Schnitte sind:

(1a) $$K^s \cdot K^r = (-1)^{(n-r)(n-s)} K^r \cdot K^s;$$

(1b) $$-K^r \cdot K^s = (-K^r) \cdot K^s = K^r \cdot (-K^s);$$

(1c) $$K^r \cdot (K^s \cdot K^t) = (K^r \cdot K^s) \cdot K^t;$$

(2) $$R(K^r \cdot K^s) = K^r \cdot R(K^s) + (-1)^{n-s} R(K^r) \cdot K^s$$

und die gewöhnlichen Distributivgesetze. Die ersten drei sind ziemlich trivialer Natur; die letzte ist die Zauberformel, die alle Existenz-, Invarianz- und Deformationssätze für Schnittpunktszahlen, Abbildungsgrade usw. in sich enthält.

Aus (2) folgt sofort:

In jedem Gebiet[14]*, das alle gemeinsamen Punkte von K^r und K^s enthält, gilt die Homologie*:

(3) $$K^r \cdot R(K^s) \sim (-1)^{n-s-1} R(K^r) \cdot K^s.$$

Ist speziell $R(K^r)$ zu K^s punktfremd, so wird die rechte Seite Null. Ersetzt man noch s durch $s+1$ und setzt man $R(K^{s+1}) = K_1^s - K_2^s$, so folgt:

Ist $K_1^s \sim K_2^s$ in einem Gebiet, das zum Rand von K^r fremd ist, so ist in diesem Gebiet auch

(4) $$K^r \cdot K_1^s \sim K^r \cdot K_2^s.$$

Sind nun zwei beliebige (nicht notwendig geradlinige) orientierte Komplexe K^r, K^s gegeben, und ist K^r zum Rand von K^s und K^s zum Rand von K^r fremd, ist weiter $\delta \leq$ dem Minimum der halben Abstände von K^r zum Rand von K^s und von K^s zum Rand von K^r, und δ-approximiert man K^r und K^s derart durch geradlinige Komplexe K_1^r und K_1^s, daß die Simplizes von K^r zu denen von K^s allgemeine Lage haben, so wird der orientierte Schnittkomplex $K_1^r \cdot K_1^s$ als ein *approximativer Schnittkomplex von K^r und K^s* bezeichnet.

---

[14]) Gebiet heißt hier eine jede offene Menge des Raumes R^n.

Inwieweit dieser approximative Schnittkomplex von der Wahl der Approximation unabhängig ist, lehrt der folgende Satz:

Ist U eine beliebige offene Umgebung der Menge aller gemeinsamen Punkte von K^r und K^s, die keine Randpunkte von K^r oder K^s enthält, und ist die Approximationsschranke $\delta \leqq$ dem Minimum der halben Abstände des außerhalb U gelegenen Teils von K^r zu K^s und des außerhalb U gelegenen Teils von K^s zu K^r, so liegt der ganze approximative Schnittkomplex $K_1^{n-r-s} = K_1^r \cdot K_1^s$ innerhalb U, und je zwei zu verschiedenen Approximationen gehörige Schnittkomplexe sind innerhalb U homolog. (In dem Extremfall, wo U den ganzen Raum mit Ausnahme der Randpunkte von K^r und K^s ausfüllt, ist die im Satz gegebene Schranke für δ dieselbe wie oben, sonst möglicherweise kleiner.)

Da der Satz mit dieser genauen Schrankenangabe nicht bei Lefschetz steht, führe ich den Beweis an: Man bilde die verbindenden Komplexe von K^s mit K_1^s und K_2^s, deren Differenz ein verbindender Komplex von K_1^s mit K_2^s ist. Wenn dieser Komplex nicht geradlinig ist, mache man ihn geradlinig durch eine Approximation, ohne die schon von vornherein geradlinigen Simplizes von K_1^s und K_2^s zu verrücken und ohne die δ-Umgebung von K^s zu verlassen. Der so konstruierte Komplex K_v^{s+1} hat (bei passender Wahl der Approximation) allgemeine Lage zu K_1^r und hat als Rand

$$R(K_v^{s+1}) = K_1^s - K_2^s - K_v^s,$$

wobei K_v^s in der δ-Umgebung von $R(K^s)$ liegt. Aus (3), angewandt auf K_1^r und K_v^{s+1}, findet man wegen $R(K^r) \cdot K_v^{s+1} = 0$:

$$K_1^r \cdot (K_1^s - K_2^s + K_v^s) \sim 0 \quad (\text{in } U),$$

oder wegen $K_1^r \cdot K_v^s = 0$:

$$K_1^r \cdot K_1^s \sim K_1^r \cdot K_2^s,$$

q. e. d.

Auf Grund der bewiesenen Eindeutigkeit bis auf Homologie bezeichnet man den approximativen Schnittkomplex $K_1^r \cdot K_1^s$ einfach mit $K^r \cdot K^s$. Mit dieser Bezeichnung gelten auch für beliebige (nicht notwendig geradlinige) Komplexe die Gleichungen (1), (3), (4) unter denselben Voraussetzungen, während (2) im betrachteten Fall nur besagt, daß $K^r \cdot K^s$ ein Zykel ist.

Wichtig ist insbesondere der Fall $r + s = n$, wo der Schnittkomplex aus endlichvielen Punkten besteht. Die eindeutige Bestimmtheit bis auf Homologie in U besagt in diesem Fall, daß bei jeder Zerlegung von U in getrennte Teilgebiete U' die algebraische Summe der Vorzeichen der Schnittpunkte in jedem Gebiet U' (kurz: die *Schnittpunktszahl in U'*) unabhängig von der Approximation ist. Auch die *Gesamtschnittpunktszahl* oder Schnittpunktzahl in U ist durch K^r und K^s allein eindeutig bestimmt: sie wird

mit $\chi(K^r, K^s)$ oder einfach mit $(K^r \cdot K^s)$ bezeichnet. Ist speziell P ein isolierter Punkt des Durchschnitts von K^r und K^s, und wählt man für ein U' eine beliebig kleine Umgebung von P, deren abgeschlossene Hülle keine weiteren Schnittpunkte enthält, so heißt die Schnittpunktszahl in U' der *Index* des Schnittpunktes P. Er kann (wie alle Schnittpunktszahlen) positiv, negativ oder Null sein. Sind alle Schnittpunkte isoliert, so ist die Indexsumme die Gesamtschnittpunktszahl.

Ist insbesondere $r = 0$, $s = n$, K^0 ein Punkt P mit positiver Orientierung, so geht die Definition des Index von P als Schnittpunkt von K^n und K^0 in die des *Abbildungsgrades* der Abbildung des Urbildkomplexes im Raum R^n im Punkt P über[15]; dieser ist eindeutig definiert, sobald P kein Randpunkt von K^n ist. Die Grundeigenschaften des Abbildungsgrades sind aus den Formeln (3), (4) abzulesen. Eine weitere Eigenschaft wird im folgenden oft gebraucht: *Der Grad einer eineindeutigen Abbildung ist* ± 1. Das folgt aus der leicht zu beweisenden Tatsache, daß der Grad der aus der Abbildung und ihrer Inversen zusammengesetzten identischen Abbildung einerseits gleich dem Produkt der beiden Abbildungsgrade, andererseits gleich $+1$ sein muß. Die Abbildungen vom Grade $+1$ sind die, welche „die Orientierung erhalten".

Nach Lefschetz[16] sind die Schnittpunktszahlen und Schnittkomplexe (bis auf Homologie) invariant bei topologischen Abbildungen dieses Gebietes mit Abbildungsgrad $+1$ (bei den anderen vom Grade -1 kehren die Schnittpunktszahlen ihr Vorzeichen um).

§ 4.

Schnittpunkte von differenzierbaren Komplexen.

Wir haben gesehen, daß eine Orientierung eines geradlinigen Simplex X^n des Raumes R^n dadurch festgelegt werden kann, daß man den Koordinaten eine bestimmte Reihenfolge zuerkennt. Dasselbe gilt nun, wenn im Simplex X^n nicht schiefwinklige, sondern beliebige (krummlinige) Koordinaten $u_1, \ldots, u_n$ eingeführt werden derart, daß das Simplex ein topologisches Bild eines Bereiches im u-Raum wird. Da nämlich der Abbildungsgrad einer eineindeutigen Abbildung stets ± 1 ist, so kann man (nach erfolgter Orientierung des u-Raums durch Wahl einer Reihenfolge $u_1, \ldots, u_n$) die Orientierung des Simplex in einer und nur einer Weise so wählen, daß der Grad der Abbildung des Simplex auf den u-Raum gleich $+1$ wird.

---

[15]) L. E. J. Brouwer, Über Abbildungen von Mannigfaltigkeiten, Math. Annalen 71 (1911), S. 97–115.

[16]) L_1 S. 21–26.

Ein Komplex K^r in R^n heißt (stückweise) *differenzierbar*, wenn für jedes Simplex des Urbildes Koordinaten $u_1, \ldots, u_r$ so eingeführt werden können, daß die Koordinaten des Bildpunktes $x_1, \ldots, x_r$ im Innern eines jeden Simplex differenzierbare, am Rande stetige Funktionen von $u_1, \ldots, u_r$ sind. Durch Wahl einer bestimmten Reihenfolge der Parameter u_j für jedes einzelne Simplex wird K^r orientiert; diese Orientierung ist noch ganz allgemein, da man sie durch Ersetzung eines u_j durch $-u_j$ für jedes einzelne Simplex in die entgegengesetzte überführen kann.

Ist P ein isolierter Schnittpunkt der beiden differenzierbaren Komplexe K^r, K^s $(r+s=n)$, ist weiter in beiden Komplexen P das Bild je eines einzelnen inneren Punktes eines Simplex mit Koordinaten $u_1, \ldots, u_r$ bzw. $v_1, \ldots, v_s$, ist schließlich an dieser Stelle die Determinante

$$D = \left| \frac{\partial x_j}{\partial u_1}, \ldots, \frac{\partial x_j}{\partial u_r}, \frac{\partial x_j}{\partial v_1}, \ldots, \frac{\partial x_j}{\partial v_s} \right| \neq 0,$$

so hat P als Schnittpunkt von K^r und K^s den Index $\varepsilon = \pm 1 = \operatorname{sign} D$.[17]

Beweis. Die x können linear so transformiert werden, daß die Determinante (im Punkt P) in der Hauptdiagonale lauter Einsen und überall sonst Nullen hat. Dabei kehrt sich die Orientierung des Raumes R^n, also auch der Index von P als Schnittpunkt, um oder nicht, je nachdem D positiv oder negativ war (§ 1). Auf Grund der Differenzierbarkeit der Funktionen $x(u)$ hat der Abstand des durch sie dargestellten Punktes vom Punkt $(u_1, \ldots, u_r, 0, \ldots, 0)$ für $|u_1| \leqq h, \ldots, |u_r| \leqq h$ die Größenordnung $o(h)$, mithin ist er bei passender Wahl von h kleiner als $\frac{h}{4}$. Ebenso haben die Punkte $x(v)$ von K^s von den Punkten $(0, \ldots, 0, v_1, \ldots, v_s)$ für $|v_j| \leqq h$ Entfernungen $< \frac{h}{4}$. Setzt man $\frac{h}{4} = \delta$, so kann man eine δ-Approximation der beiden durch $|u_j| \leqq h$, $|v_j| \leqq h$ gegebenen Ausschnitte von K^r und K^s dadurch bestimmen, daß man die Punkte x durch $(u_1, \ldots, u_r, 0, \ldots, 0)$ bzw. $(0, \ldots, 0, v_1, \ldots, v_s)$ ersetzt. Der Rand eines jeden dieser approximierenden Komplexausschnitte hat vom anderen eine Entfernung h, also hatte vor der Approximation der Rand eines jeden Ausschnittes vom anderen Ausschnitt eine Entfernung $> h - 2\delta = 2\delta$. Für die Bestimmung des Index von P als Schnittpunkt von K^r und K^s genügt es, sich auf die betrachteten Ausschnitte von K^r und K^s zu beschränken; für diese ist auf Grund der Randabstände eine δ-Approximation erlaubt (vgl. die Definition des Index in § 3); mithin hat man, um den Index von P zu bestimmen, nur das Vorzeichen von P als Schnitt-

[17]) Entsprechendes gilt für isolierte Schnittpunkte von drei oder mehr differenzierbaren Komplexen.

punkt der beiden geradlinigen Komplexe

$$'K^r:\quad (x_1, \ldots, x_n) = (u_1, \ldots, u_r, 0, \ldots, 0) \qquad (|u_j| \leqq h),$$
$$'K^s:\quad (x_1, \ldots, x_n) = (0, \ldots, 0, v_1, \ldots, v_s) \qquad (|v_j| \leqq h)$$

zu bestimmen; dieses ist aber, wie man durch eine einfache Wahl zweier kleinen Simplizes in $'K^r$ und $'K^s$ und Kombination der Ecken zu einem Simplex in R^n erkennt, gleich $+1$. Damit ist alles bewiesen.

§ 5.

Schnittpunkte von analytischen Komplexen.

Für die Untersuchung der analytischen Komplexe brauchen wir zunächst einige funktionentheoretische Begriffe und Sätze.

Wenn für alle Wertsysteme der komplexen Variablen $z_1, \ldots, z_n$ in der Umgebung einer Stelle z^0 endlichviele Wertsysteme $w_1, \ldots, w_m$ so definiert sind, daß die symmetrischen Funktionen dieser k Wertsysteme reguläre analytische Funktionen der z sind, also die w selbst Wurzeln von algebraischen Gleichungen mit analytisch von den z abhängigen Koeffizienten, so folgt daraus bekanntlich, daß zu jedem passend gewählten Wertsystem z in dieser Umgebung die w in k eindeutige analytische „Funktionszweige" zerfallen. Nimmt man nun außerdem an, daß man in beliebig kleinen Umgebungen der Ausgangsstelle aus einem dieser Funktionszweige alle anderen durch analytische Fortsetzung erhalten kann (oder daß das durch die w definierte „analytische Gebilde" nicht „zerfällt"), so soll das durch einen dieser Funktionenzweige gegebene System $w_1, \ldots, w_m$ ein *System von m mehrdeutigen analytischen Funktionen in der Umgebung der Stelle z^0* heißen. Die eindeutigen analytischen Funktionen der w und z bilden dann einen *Funktionenkörper*. Die eindeutigen analytischen Zweige aller Funktionen w sind in der betrachteten Umgebung überall regulär bis auf einem „Verzweigungsgebilde" von höchstens $2n-2$ Dimensionen im n-dimensionalen Raum der $z_1, \ldots, z_n$.

Irgend r von den Funktionen $w_1, \ldots, w_m$ heißen *analytisch-unabhängig*, wenn keine als analytische Funktion der übrigen ausgedrückt werden kann. Unter je m Funktionen w_i, die nicht alle konstant sind, sind immer eine Anzahl analytisch-unabhängiger zu finden, von denen die übrigen analytisch abhängen. Ein Kriterium für die analytische Abhängigkeit von n analytischen Funktionen von n Variablen ist das identische Verschwinden der Funktionaldeterminante.

Es gilt weiter der folgende „Umkehrsatz": *Ist w eine mehrdeutige analytische Funktion von $z_1, \ldots, z_n$ in der Umgebung einer Stelle z^0 in einem Funktionenkörper und ist $\frac{\partial w}{\partial z_1}$ nicht identisch Null, so können um-*

gekehrt z_1 und daher weiter alle Funktionen des Körpers auch als mehrdeutige analytische Funktionen von w, $z_2, \ldots, z_n$ in einer Umgebung der Stelle w^0, $z_2^0, \ldots$ aufgefaßt werden. (Ist dagegen $\frac{\partial w}{\partial z_1} = 0$, so ist w eine Funktion von $z_2, \ldots, z_n$ allein).

Durch wiederholte Anwendung folgt der „Austauschsatz“: *Sind $w_1, \ldots, w_m$ analytisch-unabhängige analytische Funktionen von $z_1, \ldots, z_n$ in einem Funktionenkörper, so gibt es unter den z_j m Größen, die gegen die w ausgetauscht werden können, derart, daß alle Funktionen des Körpers als analytische Funktionen der w und der übrigbleibenden z aufgefaßt werden können.*

Ist insbesondere $m = n$, so folgt, daß alle z_j mehrdeutige analytische Funktionen der w in der Umgebung der Stelle w^0 werden.

n komplexe Veränderliche $z_j = x_j + i y_j$ können als komplexe Koordinaten in einem R^{2n} aufgefaßt werden. Für die reellen Koordinaten x_j, y_j setzen wir ein für allemal die Reihenfolge

$$x_1, y_1, x_2, y_2, \ldots, x_n, y_n$$

fest, wodurch nach § 2 eine bestimmte Orientierung des Raumes R^{2n} gegeben ist.

Ein Komplex K^{2r} im R^{2r}, der aus eineindeutig stetigen Bildern von geradlinigen (2)-dimensionalen Simplizes aufgebaut ist, heißt ein *analytischer Komplex von r komplexen Dimensionen*, wenn die Umgebung einer beliebigen Stelle von K^{2r} (Randstellen ausgenommen) sich aus endlichvielen „analytischen Zweigen“ zusammensetzt, auf denen jeweils die Koordinaten $z_1, \ldots, z_n$ als mehrdeutige analytische Funktionen von r unter ihnen, etwa von $z_1, \ldots, z_r$, aufgefaßt werden können.

Diese Struktur der Umgebungen einer Stelle besitzen nicht nur die algebraischen Varietäten, sondern nach O. Blumenthal[18]) überhaupt alle durch analytische Gleichungen definierbaren Punktmengen. Sollen diese unter unsere Definition fallen, so müssen sie außerdem Komplexe, im obigen Sinn, sein, d. h. mit endlichvielen eineindeutig stetigen Bildern von Simplizes triangulierbar sein. Diese Bedingung kann man dadurch erfüllen, daß man sich auf einen beschränkten Bereich $|z_j| \leqq M$ beschränkt, in dem die definierenden Gleichungen des untersuchten Gebildes überall regulär sind (vgl. Anhang 1). Richtet man die Triangulation so ein, daß die Verzweigungsgebilde der analytischen Funktionen z_j immer aus höchstens $(2r-2)$-dimensionalen Seiten der zur Triangulation benutzten Simplizes bestehen, so kann man in den einzelnen Simplizes die unabhängigen

[18]) O. Blumenthal, Zum Eliminationsproblem bei analytischen Funktionen mehrerer Veränderlicher, Math. Annalen 57 (1903), S. 356.

$z_1, \ldots, z_r$ als Koordinaten benutzen. Die übrigen Raumkoordinaten sind dann im Innern dieser Simplizes differenzierbare Funktionen dieser $2r$ Parameter, mithin:

Ein analytischer Komplex von r komplexen Dimensionen ist ein differenzierbarer Komplex von $2r$ Dimensionen.

Statt der speziellen Parameter $z_1, \ldots, z_r$ benutzen wir im folgenden beliebige komplexe Parameter $w_1, \ldots, w_r$, von denen die Koordinaten $z_1, \ldots, z_n$ im Innern eines Simplex regulär-analytisch abhängen. Setzt man $w_j = u_j + i v_j$, so ist durch die zu Anfang dieses Paragraphen gemachten Voraussetzungen über die Reihenfolge der Real- und Imaginärteile eine bestimmte Orientierung aller Simplizes des Komplexes K^{2r} festgelegt, unabhängig von der Reihenfolge der w.

Beim Übergang zu anderen Koordinaten $W_1, \ldots, W_r$, die von den w umkehrbar-analytisch abhängen, ist $\frac{\partial(W_1, \ldots, W_r)}{\partial(w_1, \ldots, w_r)} \neq 0$. Wir wollen, nach Aufspaltung in Real- und Imaginärteil $W_j = U_j + i V_j$, die Funktionaldeterminante $\frac{\partial(U_1, V_1, \ldots, U_r, V_r)}{\partial(u_1, v_1, \ldots, u_r, v_r)}$ berechnen. Sie ändert ihren Wert nicht, wenn wir in Zähler und Nenner formal Linearkombinationen $W = U + iV$, $\overline{W} = U - iV$ einführen. Also hat man

$$\frac{\partial(U_1, V_1, \ldots, U_r, V_r)}{\partial(u_1, v_1, \ldots, u_r, v_r)} = \frac{\partial(W_1, \overline{W}_1, \ldots, W_r, \overline{W}_r)}{\partial(w_1, \overline{w}_1, \ldots, w_r, \overline{w}_r)}$$

$$= \frac{\partial(W_1, \ldots, W_r)}{\partial(w_1, \ldots, w_r)} \cdot \frac{\partial(\overline{W}_1, \ldots, \overline{W}_r)}{\partial(\overline{w}_1, \ldots, \overline{w}_r)} = \left| \frac{\partial(W_1, \ldots, W_r)}{\partial(w_1, \ldots, w_r)} \right|^2 > 0 .$$

Daraus folgt, daß die oben festgesetzte Orientierung sich bei Übergang zu den Parametern W nicht ändert.

Die Parameter in zwei längs einer $(2r-1)$-dimensionalen Seite aneinanderstoßenden Simplizes von K^{2r} hängen auf dieser Seite durch eine analytische Transformation miteinander zusammen, denn die Singularitäten der Parameterdarstellung liegen alle auf den höchstens $(2r-2)$-dimensionalen Seiten. Die Orientierungen in diesen anstoßenden Simplizes stimmen also überein, d. h. *der analytische Komplex K^{2r} ist orientierbar.*

Die Randpunkte analytischer Komplexe, wo die analytische Darstellung eventuell aufhören kann, mögen ein für allemal von der Untersuchung ausgeschlossen werden.

Nunmehr seien zwei analytische Komplexe K^{2r}, K^{2s} $(r+s=n)$ in der Umgebung eines isolierten (inneren) Schnittpunktes P gegeben durch Parameterfunktionen $z_j(w_1, \ldots, w_r)$ bzw. $z_j^*(w_1^*, \ldots, w_s^*)$. Im einfachsten Fall, daß P ein innerer Punkt eines Simplex von K^{2r} und eines Simplex von K^{2s} ist, und die komplexe Funktionaldeterminante $\left| \frac{\partial z_j}{\partial w_1}, \ldots, \frac{\partial z_j}{\partial w_r}, \right.$

$\left.\frac{\partial z_j^*}{\partial w_1^*}, \ldots, \frac{\partial z_j^*}{\partial w_s^*}\right|$ in $P \neq 0$ ist, findet man durch dieselbe Rechnung wie oben für die reelle Determinante D von § 4 einen positiven Wert; *mithin hat in diesem Fall der Schnittpunktsindex den Wert* $+1$.

Wir wollen nun für einen beliebigen isolierten Schnittpunkt beweisen, daß der Schnittpunktsindex positiv sein muß, und zwar wollen wir durch eine kleine Parallelverschiebung den Fall auf den vorigen zurückführen.

Wir setzen zunächst $t_j = z_j - z_j^*$, und wollen zeigen, daß die n Funktionen t_j der n Veränderlichen w, w^* analytisch-unabhängig sind. Wären sie es nicht, und wären etwa nur $t_1, \ldots, t_m$ $(m < n)$ analytisch-unabhängig, so könnte man nach dem Austauschsatz $n - m$ von den Größen w und w^* so auswählen, daß alle z, z^*, w, w^* als mehrdeutige analytische Funktionen von $t_1, \ldots, t_m$ und den $n - m$ übriggebliebenen w, w^* aufgefaßt werden können. Läßt man nun, von den zum Schnittpunkt gehörigen Nullwerten dieser neuen Variablen ausgehend, ein w oder w^* sich stetig ändern, während man $t_1, \ldots, t_m$ konstant gleich Null läßt (man kann sich dabei ganz beliebig für irgendeinen Zweig der mehrdeutigen Funktionen z, z^*, w, w^* entscheiden), so bleiben auch die übrigen t_j, die ja analytisch von $t_1, \ldots, t_m$ abhängen, konstant gleich Null, also $z_j = z_j^*$ $(j = 1, \ldots, n)$, mithin hat man es fortwährend mit einem Schnittpunkt zu tun, während doch ein w oder w^* variabel ist, entgegen der Voraussetzung, wir hätten es mit einem isolierten Schnittpunkt zu tun.

Also sind die t_j analytisch unabhängig. Daraus folgt erstens, daß ihre Funktionaldeterminante $D = \frac{\partial(t_1, \ldots, t_n)}{\partial(w_1, \ldots, w_s^*)}$ nicht identisch verschwindet, und zweitens, daß die w, w^* und z, z^* als endlichvieldeutige analytische Funktionen der t_j in der Umgebung der Null aufgefaßt werden können. Zu jeder Stelle t in dieser Umgebung gehören endlichviele Stellenpaare z, z^* mit $z_j - z_j^* = t_j$, und bei passend gewählten t (nämlich überall bis auf einem Gebilde von höchstens $2n - 2$ Dimensionen) sind an allen diesen Stellen die Funktionen $z(w)$, $z^*(w^*)$ unverzweigt und die Funktionaldeterminante D nicht Null.

Geometrisch bedeutet das: Wenn man einen der beiden gegebenen Komplexe, etwa K^{2r}, um beliebig wenig parallel verschiebt, indem man den Punkt z_j durch $z_j - t_j$ ersetzt, so hat der verschobene Komplex K^{2r} mit dem durch $z^*(w^*)$ gegebenen K^{2s} endlichviele Punkte in der Umgebung der betreffenden Stelle gemein, und bei passender Wahl der t_j sind alle diese Stellen innere Punkte der Simplizes von K^{2r} und K^{2s}, während die Funktionaldeterminante

$$D = \frac{\partial(z_1 - z_1^*, \ldots, z_n - z_n^*)}{\partial(w_1, \ldots, w_r, \ldots, w_s^*)} = (-1)^s \cdot \left| \frac{\partial z_j}{\partial w_1}, \ldots, \frac{\partial z_j}{\partial w_r}, \frac{\partial z_j^*}{\partial w_1^*}, \ldots, \frac{\partial z_j^*}{\partial w_s^*} \right|$$

an diesen Stellen nicht verschwindet.

Daraus folgt nach dem früheren, daß diese endlichvielen Schnittpunkte der verschobenen Komplexe die Indizes $+1$ haben. *Somit hat der ursprüngliche isolierte Schnittpunkt einen positiven Index, gleich der Anzahl der Schnittpunkte nach der Verschiebung.*

Der obige Beweis ist, bis auf die Ersetzung der algebraischen Abhängigkeiten durch analytische und einer kleinen Vereinfachung dadurch, daß eine Verschiebung statt einer projektiven Transformation benutzt wurde, wesentlich derselbe wie der Beweis für die Positivität der Schnittpunktsmultiplizitäten in W_2, § 9. Zugleich stellt er die Ausführung der Lefschetzschen Beweisandeutung (Fußnote [12])) dar. Er ist mühelos auf Schnittpunkte von mehr als zwei Komplexe auszudehnen.

§ 6.

Die Beziehung zum algebraischen Multiplizitätsbegriff.

Dieser Paragraph ist im logischen Zusammenhang dieser Arbeit entbehrlich und dient bloß dazu, den Anschluß der jetzigen Untersuchung an die frühere W_2 herzustellen.

Die durch algebraische Gleichungen definierten Punktmengen oder *algebraischen Varietäten* im komplexen projektiven Raum sind triangulierbare Komplexe (vgl. Anhang 1), die in irreduzible Bestandteile von verschiedener Dimension zerfallen, welche in der Umgebung einer jeden Stelle analytisch sind. Wir bezeichnen sie mit V_r oder V^{2r}, wenn r die algebraische oder komplexe Dimension und daher $2r$ die reelle Dimensionszahl ist. Orientierung nach der Vorschrift von § 5.

Die *algebraische Multiplizität* eines Schnittpunktes P zweier algebraischen Varietäten V_r, V_s $(r+s=n)$ kann folgendermaßen definiert werden (vgl. W_2, § 8): Man bringe zunächst die beiden Varietäten in „allgemeine Lage" zueinander, indem man die eine, etwa V_r, einer projektiven Transformation mit unbestimmten Koeffizienten unterzieht. Die Schnittpunkte der transformierten Varietäten (jeder einmal gezählt) ergeben, wenn man nachher die Transformationsmatrix zur Einheitsmatrix spezialisiert, durch „relationstreue Spezialisierung" ebenso viele (gleiche oder verschiedene) Schnittpunkte der untransformierten Varietäten, und die Zahl, die angibt, wie oft dabei ein jeder Schnittpunkt entsteht, ist die „Multiplizität" dieses Schnittpunkts. Dabei wird unter einer „relationstreuen Spezialisierung" eine solche verstanden, wobei alle durch homogene Gleichungen ausgedrückten algebraischen Relationen, die zwischen den Schnittpunkten und den Transformationskoeffizienten bestehen, erhalten bleiben (vgl. W_1, § 3).

Wählt man die uneigentliche Ebene des Raumes so, daß weder vor noch nach der Transformation Schnittpunkte in ihr liegen (das geht, wenn,

wie wir annehmen wollen, nur endlichviele Schnittpunkte vorhanden sind), so kann man diese uneigentliche Ebene ganz weglassen, d. h. sich auf den euklidischen Raum beschränken, und die homogenen Gleichungen, die in der Definition der Relationstreue vorkamen, durch inhomogene ersetzen. Weiter kann man die Matrix mit unbestimmten Koeffizienten ersetzen durch eine von der Einheitsmatrix beliebig wenig verschiedene Matrix, deren Elemente algebraisch-unabhängige Transzendente sind (unabhängig auch von den Koeffizienten der definierenden Gleichungen der gegebenen Varietäten); für die algebraischen Eigenschaften der Schnittpunkte macht das offenbar nichts aus. Schließlich kann man die relationstreue Spezialisierung durch einen Grenzübergang ersetzen, denn bei einem solchen bleiben tatsächlich alle durch algebraische Gleichungen ausdrückbaren Relationen erhalten.

Bei einem Grenzübergang bleiben aber auch alle topologischen Schnittpunktszahlen im Sinne von § 3 [19]) erhalten, denn diese Anzahlen waren durch eine Approximation definiert, welche natürlich für alle diejenigen transformierten Varietäten, welche sich von der gegebenen um hinreichend wenig unterscheiden, gültig bleibt. Also ist die Schnittpunktsanzahl in der Umgebung einer Stelle P dieselbe für eine der approximierenden V_r mit der festen V_s wie für die gegebene V_r mit derselben V_s.

Wie wir wissen, können wir in beliebiger Nähe von V_r parallel-verschobene V_r finden, für die alle Schnittpunktsindizes mit V_s gleich Eins werden. Die Matrix einer solchen Parallelverschiebung (als projektive Transformation aufgefaßt) sei B, und eine Matrix mit unabhängigen transzendenten Koeffizienten, die sich von B sehr wenig unterscheidet, sei A. Wir können dann den Grenzübergang, der unsere relationstreue Spezialisierung liefern soll, so ausführen, daß wir zuerst A zu B, dann B zur Einheitsmatrix E konvergieren lassen. Würden beim ersten Grenzübergang $A \to B$ mehrere Schnittpunkte zu einem konvergieren, so würde der Index dieses einen Grenzpunktes größer als Eins werden, was unmöglich ist. Also konvergieren bei $A \to B$ alle Schnittpunkte wieder zu getrennten Schnittpunkten. Beim Übergang $B \to E$ rücken einige Schnittpunkte in P zusammen; ihre Anzahl ist erstens gleich dem Index von P (da die Indexsumme in der Umgebung erhalten bleibt), zweitens gleich der algebraischen Multiplizität (da es sich um eine relationstreue Spezialisierung handelt). Damit ist bewiesen:

Die algebraische Multiplizität eines Schnittpunktes zweier algebraischen Mannigfaltigkeiten, die nur endlichviele Schnittpunkte besitzen, ist gleich dem topologischen Index.

---

[19]) Um diese Definition anwenden zu können, beschränke man sich jeweils auf einen ganz im Endlichen gelegenen Teil der Komplexe V_r bzw. V_s.

Die in W_2, § 9, auf algebraischem Wege bewiesene Tatsache, daß die Multiplizität eines jeden isolierten Schnittpunktes stets größer als Null ist, ergibt sich jetzt als Spezialfall des allgemeineren Ergebnisses von § 5 für analytische Komplexe.

§ 7.

Ausdehnung auf allgemeinere Mannigfaltigkeiten.

Unter einer *topologischen Mannigfaltigkeit* wird verstanden ein zusammenhängender topologischer Raum, in dem eine Umgebung eines jeden Punktes topologisch auf eine offene Menge eines euklidischen Parameterraums abbildbar ist. Ist der Raum kompakt (etwa ein Komplex), so spricht man von einer *geschlossenen Mannigfaltigkeit*. An jeder Stelle, wo zwei verschiedene euklidisch-abbildbare Umgebungen sich teilweise überdecken, hängen die beiden Parameterdarstellungen durch eine eineindeutige Parametertransformation mit Abbildungsgrad ± 1 miteinander zusammen. Können die Orientierungen der euklidischen Parameterräume so gewählt werden, daß diese Abbildungsgrade alle gleich $+1$ werden, so heißt die Mannigfaltigkeit *orientierbar*. Sind die genannten Parametertransformationen beiderseits differenzierbar, so spricht man von einer *differenzierbaren Mannigfaltigkeit*; auf ihr kann man in naheliegender Weise differenzierbare und nicht-differenzierbare Komplexe unterscheiden. Sind die Parametertransformationen sogar (nach Zusammenfassung je zweier reeller Parameter x, y zu einem komplexen $x + iy$) analytisch, so spricht man von einer (komplexen) *analytischen Mannigfaltigkeit* (vgl. die Weylsche Definition einer Riemannschen Fläche); eine solche ist von selbst orientierbar, da der Abbildungsgrad einer analytischen Abbildung im komplexen Gebiet stets positiv ist (§ 5).

Es sei M^n eine orientierbare und orientierte topologische Mannigfaltigkeit. Wir betrachten zwei Komplexe K^r, K^s $(r + s = n)$ in M^n, bei denen der Rand von K^r zu K^s und der Rand von K^s zu K^r fremd ist. Der *Index* eines isolierten Schnittpunkts P von K^r und K^s wird definiert, indem die Umgebung von P vermöge einer der Parameterdarstellungen von M^n auf eine offene Menge des euklidischen R^n abgebildet wird und dort der Index im Sinne von § 2 bestimmt wird. Diese Indexdefinition ist invariant gegenüber Parametertransformationen mit Abbildungsgrad $+1$. Sind alle Schnittpunkte von K^r und K^s isoliert, so heißt die Indexsumme die *Schnittpunktszahl* von K^r und K^s.

Zur Definition der Schnittpunktszahlen und Schnittkomplexe bei nichtisolierten Schnittpunkten und für $r + s > n$ braucht man simpliziale Ap-

proximationen, die nicht nur für die Umgebung einer Stelle, sondern in der ganzen Mannigfaltigkeit sinnvoll bleiben, und zwar braucht man verschiedene Approximationen für K^r und K^s derart, daß eine Zelle der Approximation von K^r mit einer Zelle der Approximation von K^s immer einen Schnittkomplex von der richtigen Dimension $r+s-n$ hat. Die Möglichkeit einer solchen Approximation ist bis heute nur bewiesen für *„Mannigfaltigkeiten im kombinatorischen Sinn“*, die derart in Simplizes eingeteilt werden können, daß die Simplizes um einen Punkt sich so aneinanderschließen wie Simplizes im euklidischen Raum, die eine Umgebung einer gemeinsamen Ecke ganz ausfüllen. Wir können dafür auf die schon mehrfach genannte Lefschetzsche Arbeit (L_1) verweisen. Für allgemeinere Mannigfaltigkeiten steht der Beweis noch aus (sogar wenn sie triangulierbar sind); wir setzen daher einstweilen voraus, daß wir es mit Mannigfaltigkeiten im kombinatorischen Sinn zu tun haben[20]).

Die Formeln (1), (3), (4) von § 3 können für beliebige Komplexe genau so bewiesen werden wie für den euklidischen Raum. Auch die Invarianz bei topologischen Abbildungen vom Grade $+1$ hat Lefschetz bewiesen; daraus folgt insbesondere die Übereinstimmung des mit diesen Methoden definierten Indexbegriffs mit dem vorhin durch direkte Übertragung aus dem euklidischen Raum definierten Indexbegriff im Fall eines isolierten Schnittpunkts.

Für analytische Komplexe in einer analytischen Mannigfaltigkeit sind, wenn man alle Orientierungen gemäß der Verabredung in § 5 *normiert, die Indizes isolierter Schnittpunkte stets positiv*[21]). Das folgt unmittelbar aus dem Satz von § 5.

Da außerdem die Dimensionen analytischer Komplexe gerade sind, gilt im kommutativen Gesetz (1a) für sie stets das $+$-Zeichen.

---

[20]) **Zusatz bei der Korrektur.** Inzwischen ist die Dissertation von E. R. van Kampen („Die kombinatorische Topologie und die Dualitätssätze“, Leiden 1929) erschienen, in der die erwähnte Lücke ausgefüllt ist. Dort werden nämlich die Schnittpunktszahlen definiert und ihre Eigenschaften bewiesen für triangulierbare „Mannigfaltigkeiten“ in einem sehr allgemeinen Sinn, unter die unter anderem alle oben definierten „topologischen Mannigfaltigkeiten“ fallen, soweit sie triangulierbar sind. Für die algebraische Anwendung reicht das völlig hin, denn alle algebraischen Mannigfaltigkeiten sind triangulierbar (Anhang 1).

[21]) Bei nichtisolierten Schnittpunkten können aber die Schnittpunktszahlen sehr wohl negativ sein, wie das folgende Beispiel zeigt: Durch einen Kegelschnitt K im komplexen projektiven R_3 werde eine singularitätenfreie Fläche 5. Ordnung F gelegt. Die Ebene des Kegelschnittes schneidet die Fläche außerdem in einer kubischen Kurve C. Die Schnittpunktszahl von K mit jedem ebenen Schnitt von F auf F, also insbesondere mit $K+C$, ist $+2$, die von K mit C aber $+6$, also bleibt für die Schnittpunktszahl von K mit K nur -4 übrig.

Ist die Mannigfaltigkeit, in der wir uns befinden, der komplexe projektive Raum oder allgemeiner eine (singularitätenfreie) allgemeine Varietät V_n,[22]) so sind alle in dieser Mannigfaltigkeit liegenden algebraischen Varietäten $V_r = V^{2r}$ erstens *triangulierbar* (vgl. Anhang 1), zweitens *orientierbar* (wie alle analytischen Komplexe) und drittens *unberandet* (weil die algebraischen Funktionen von r komplexen Variablen keine Singularitätengebilde von mehr als $2r-2$ Dimensionen haben, mithin ein $(2r-1)$-dimensionaler Rand nicht auftreten kann), also *Zyklen*.

Für isolierte Schnittpunkte algebraischer V_r und V_s in einer Mannigfaltigkeit V_n $(r+s=n)$ ist noch keine allgemein anwendbare algebraische Multiplizitätsdefinition bekannt. Wir definieren daher (in Analogie zum projektiven Raum) als *Schnittpunktsmultiplizität* den topologischen Index des betreffenden Schnittpunktes.

Die Indexsumme genügt dem „Prinzip der Erhaltung der Anzahl", nicht nur bei stetigen Änderungen der schneidenden Varietäten, sondern sogar bei Ersetzung einer V_r durch einen dazu homologen Zykel. Nennt man zwei Zyklen *homolog mit erlaubter Division* (Zeichen $\approx$), wenn ein ganzzahliges Vielfaches ihrer Differenz homolog Null ist, so ist sofort einzusehen, daß man auch einen Zykel C^k durch einen anderen $D^k \approx C^k$ ersetzen kann, ohne seine Schnittpunktszahl mit irgendeinem Zykel der komplementären Dimension zu ändern. Denn eine Schnittpunktszahl, von der ein Vielfaches Null ist, ist selbst Null.

Jetzt sei allgemeiner $r+s \geq n$. Die Durchschnittsmenge V einer V_r und einer V_s ist eine algebraische Varietät, von der wir annehmen wollen, sie habe nicht mehr als $k=r+s-n$ komplexe Dimensionen. Der approximative Schnittkomplex S^{2k} von V_r und V_s kann in einer beliebig dünnen Umgebung von V gewählt werden. Legt man nun eine Triangulation des Raumes zugrunde, bei der V aus Seiten der Triangulation besteht, und approximiert man S^{2k} mit dieser Triangulation, so wird, wenn S^{2k} in einer genügend kleinen Umgebung von V liegt, der approximierende Zykel S_1^{2k} in V liegen, d. h. aus Zellen der höchsten Dimension $2k$ von V bestehen. Sind $V', V'', \ldots$ die $2k$-dimensionalen irreduziblen Bestandteile von V, so folgt:

$$(5) \qquad V_r \cdot V_s = S^{2k} \sim S_1^{2k} = \mu' V' + \mu'' V'' + \ldots.$$

Die Größen $\mu', \mu'', \ldots$ nennen wir die *Multiplizitäten* der Schnittbestandteile $V', V'', \ldots$.[23])

[22]) Der untere Index n soll die algebraische oder analytische Dimensionszahl darstellen; die topologische Dimensionszahl ist doppelt so groß und wird als oberer Index benutzt.

[23]) Algebraisch kann man auch diese Multiplizitäten noch definieren im Fall des projektiven Raums, indem man die Varietäten V_r, V_s, V mit einem allgemeinen linearen Raum L_{n-k} schneidet und dort die Schnittpunktsmultiplizitäten bestimmt.

Bemerkung. Wenn ein irreduzibler Bestandteil von V eine Dimension $< 2k$ haben sollte, so kann er in (5) rechts nicht vorkommen. Man zeigt aber leicht, daß es solche Bestandteile gar nicht geben kann. Wäre nämlich P ein Punkt eines solchen, der nicht zugleich zu einem der $2k$-dimensionalen Bestandteile $V', V'', \ldots$ gehört, so würde man V_r, V_s und V mit einem V_{n-k} durch P schneiden können, der in P einen isolierten Schnittpunkt mit V hätte. Die Schnittpunktszahl von $V_r \cdot V_s \cdot V_{n-k}$ in einer Umgebung von P müßte nach § 5 positiv sein, andererseits aber sich nicht ändern, wenn man für $V_r \cdot V_s$ die Approximation $\mu' V' + \mu'' V'' + \ldots$ einsetzt, also (da P nicht zu $\mu' V' + \ldots$ gehört) Null sein, was ein Widerspruch ist. Die hiermit bewiesene Tatsache, daß ein Schnittgebilde von analytischen Varietäten von r und s komplexen Dimensionen keine Komponenten von weniger als $k = r + s - n$ komplexen Dimensionen besitzen kann, ist eine Verallgemeinerung einer von Blumenthal[24]) bewiesenen Behauptung, die sich auf den Fall $s = n - 1$, $k = r - 1$ bezieht.

Hat die Durchschnittsvarietät V mehr als k komplexe Dimensionen, so versagt die Darstellung (5) von $V_r \cdot V_s$ als Summe von algebraischen Varietäten und man bleibt auf die approximativen Schnittkomplexe angewiesen[25]).

§ 8.

Der Schubertsche Kalkül.

Dem Schubertschen Bedingungskalkül[26]) liegt immer eine feste singularitätenfreie algebraische Varietät $M^{2n} = M_n$ zugrunde. Die Elemente von M_n können irgendwelche geometrische Gebilde sein (etwa alle Geraden des Raumes), aber wir verlangen, daß man M_n auch als Varietät von Punkten in einem mehrdimensionalen projektiven Raum auffassen kann. Wir verlangen außerdem, daß M^{2n} eine topologische Mannigfaltigkeit ist, in welcher Schnittkomplexe durch simpliziale Approximationen definiert werden können (§ 7).

Schubert verwendet Symbole für algebraische Bedingungen, die den Punkten von M_n aufgelegt werden: eine k-fache Bedingung ist eine solche, die eine algebraische Varietät V_{n-k} in M_n definiert, deren Punkte eben dieser Bedingung genügen. Es ist bequem, für diese Bedingung dasselbe Symbol zu benutzen wie für die Varietät V_{n-k} selber.

---

[24]) O. Blumenthal, loc. cit. [19]).

[25]) Man kann beweisen, daß man auch in diesem Fall noch $V_r \cdot V_s$ als Summe oder Differenz von algebraischen Varietäten darstellen kann. Für unsere Zwecke ist das aber nicht nötig.

[26]) H. Schubert, Kalkül der abzählenden Geometrie, Leipzig 1879.

Eine „symbolische Gleichung“ zwischen Symbolen von k-fachen Bedingungen $U, V, \ldots$, etwa

(6) $$U \doteq V + \frac{1}{2} W \text{ [27]}$$

bedeutet nach Schubert die Aussage, daß für *jede* algebraische V_k der M_n die Anzahlen der Punkte von V_k, welche jeweils den Bedingungen U, V, W genügen, in derselben linearen Relation zueinander stehen wie die Symbole U, V, W in (6). Dabei sind natürlich die „Anzahlen“ zu Indexsummen, also zu Schnittpunktszahlen χ zu präzisieren. In unseren Bezeichnungen würde die symbolische Gleichung (6) also heißen, daß für jede algebraische V_k in M_n die Gleichung

(7) $$\chi(U, V_k) = \chi(V, V_k) + \frac{1}{2}(W, V_k)$$

besteht. Es ist dabei bequem, die Gleichung (7) zu verlangen nicht nur für alle algebraischen V_k, sondern auch für die k-dimensionalen approximativen Schnittkomplexe von algebraischen Varietäten höherer Dimension[28]).

In der Form (7) hat die Gleichung (6) einen ganz präzisen Sinn, der außerdem unabhängig davon ist, ob die Varietäten U und V_k, V und V_k usw. wirklich nur endlichviele Schnittpunkte haben.

Wir lassen mit Schubert in Ausdrücken wie (7) fortan die χ und die Klammer weg, verwenden also ein und dasselbe Symbol für den nulldimensionalen Komplex $U \cdot V_k$ und für die Anzahl (genauer: Summe der Vorzeichen) seiner Punkte.

Klar ist, daß man symbolische Gleichungen addieren, subtrahieren und mit rationalen Zahlen multiplizieren darf. Man darf aber auch alle Glieder mit dem Symbol einer Bedingung H multiplizieren (multiplizieren in der topologischen Bedeutung von Schnittbildung, welche in der Tat die richtige Präzisierung der Schubertschen Multiplikation von Bedingungen darstellt).

Ist nämlich H eine h-fache Bedingung, mit der man die Gleichung (6) zu multiplizieren wünscht, so kann man zunächst $h + k \geqq n$ annehmen, da sonst alle Produkte $U \cdot H$ usw. Null werden. Die Behauptung

$$U \cdot H \doteq V \cdot H + \frac{1}{2} W \cdot H$$

bedeutet nun, daß für jede algebraische V_{h+k} gilt

$$U \cdot H \cdot V_{h+k} = V \cdot H \cdot V_{h+k} + \frac{1}{2} W \cdot H \cdot V_{h+k}.$$

[27]) Schubert schreibt $=$, wo hier $\doteq$ steht.

[28]) Nach dem in Fußnote [25]) angeführten Satz macht es keinen wirklichen Unterschied, ob man dies verlangt oder nicht.

Auf Grund des Assoziativgesetzes (1c) erhält man aber dieselbe Gleichung, wenn man in (7) speziell $V_k = H \cdot V_{h+k}$ setzt. Damit ist die Behauptung bewiesen.

Man darf also auch symbolische Gleichungen miteinander multiplizieren. Für die Multiplikation gelten nach (1) die Assoziativ- und Kommutativgesetze. Damit ist der formale Teil des Bedingungskalküls exakt gerechtfertigt.

Wichtig für die Anwendungen ist insbesondere, daß *jede Homologie mit erlaubter Division zwischen algebraischen Varietäten zu einer symbolischen Gleichung im Schubertschen Sinne Anlaß gibt.* Ist nämlich $U \approx V$, so ist nach § 7 für jede V_k:

$$U \cdot V_k = V \cdot V_k,$$

mithin

$$U \doteq V.$$

Ob auch die Umkehrung gilt, ist mir nicht bekannt, aber ohne viel Interesse, denn man findet in praxi aus Stetigkeitsbetrachtungen immer zunächst Homologien, aus denen man dann sofort symbolische Gleichungen folgern kann.

Das *Charakteristikenproblem* für eine Mannigfaltigkeit M_n lautet in der Schubertschen Fassung: Es ist eine Basis $V^{(1)}, \ldots, V^{(m)}$ für die k-fachen Bedingungen aufzustellen derart, daß jede k-fache Bedingung V symbolisch gleich einer Linearkombination der Basiselemente wird:

$$V \doteq \lambda_1 V^{(1)} + \ldots + \lambda_m V^{(m)}. \tag{8}$$

Die Lösung des Charakteristikenproblems konnte bis jetzt immer nur von Fall zu Fall gegeben werden. Die topologische Methode liefert zum erstenmal einen allgemeinen Beweis für die Existenz einer Basis. Denn es gibt in jeder triangulierbaren M^{2n} eine endliche Basis für die Klassen homologer Zyklen, also um so mehr eine Basis für die Homologieklassen mit erlaubter Division. Sucht man nun aus den letzteren Klassen diejenigen aus, die algebraische Varietäten enthalten, so haben diese wiederum eine endliche Basis. Die so erhaltene Basis in bezug auf die Relation $\approx$ ist um so mehr eine Basis für die Relation $\doteq$.

Die Anzahlen der nötigen Basiselemente für die Dimensionen k und $n-k$ sind gleich, und die Matrix der Schnittpunktszahlen ist regulär. Denn sonst würde es eine Linearkombination der Basiselemente für die Dimension k oder $n-k$ geben, welche mit jeder algebraischen Varietät von der komplementären Dimension die Schnittpunktszahl Null hätte, also symbolisch gleich Null wäre; damit wäre aber ein Basiselement überflüssig. — Folglich kann man die Basis für die Dimensionen k und $n-k$ auch so wählen, daß die Matrix eine Einheitsmatrix wird. Die Koeffizienten in (8)

lassen sich dann deuten als Schnittpunktszahlen von V mit den Basiselementen der komplementären Dimension $n-k$.

Die wirkliche Aufstellung einer Homologiebasis ist in vielen Fällen nicht schwer. Hat man einmal die Basis für die Dimensionen k und $n-k$, sowie die Schnittpunktszahlen für die Basiselemente gefunden, so ist man im Stande, für je zwei (als Linearkombinationen der Basiselemente geschriebene) Varietäten V_k und V_{n-k} die Schnittpunktszahlen auszurechnen.

Für den Fall des projektiven Raumes soll das im nächsten Paragraphen ausgeführt werden.

§ 9.

Anwendung auf den projektiven Raum.

Im Anhang 2 wird bewiesen, daß die linearen Räume L_j eine Homologiebasis für alle Dimensionen im komplexen projektiven Raum $P_n = P^{2n}$ bilden. Führen wir mit Schubert für die einfache lineare Bedingung L_{n-1} das Symbol p ein, so stellt die Potenz p^h den L_{n-h} dar. Jede Varietät V_r ist homolog einem Vielfachen von L_r, also haben wir:

(9) $$V_r \doteq \alpha \cdot p^{n-r}.$$

Daß hier α den Grad von V_r (d. h. die Schnittpunktszahl mit einem L_{n-r}) darstellt, sieht man sofort, indem man die Gleichung (9) mit p^r multipliziert.

Ist für eine V_{n-r} ebenso

(10) $$V_{n-r} = \beta \cdot p^r,$$

so erhält man durch Multiplikation von (9) mit (10) sofort den *„verallgemeinerten Bézoutschen Satz"*:

(11) $$V_r \cdot V_{n-r} = \alpha\beta.$$

Sind allgemeiner V_r und V_s mit $r+s \geqq n$ gegeben, und sind α und β wie vorhin die Gradzahlen, so erhält man ebenso:

$$V_r \cdot V_s \doteq \alpha\beta \cdot p^{2n-r-s},$$

d. h. der approximative Schnittkomplex von V^r und V^s hat als Grad das Produkt der Gradzahlen von V^r und V^s. Hat der wirkliche Durchschnitt von V^r und V^s die richtige Dimension $k = r+s-n$, und zerfällt er in irreduzible Komponenten $V', V'', \ldots$ von den Gradzahlen $g', g'', \ldots$ und Multiplizitäten $\mu', \mu'', \ldots$, so hat man

$$V_r \cdot V_s = \mu' V' + \mu'' V'' + \ldots = \mu' g' p^{n-k} + \mu'' g'' p^{n-k} + \ldots,$$

mithin

(12) $$\mu' g' + \mu'' g'' + \ldots = \alpha \cdot \beta.$$

Anhang 1.

Triangulierbarkeit der algebraischen Gebilde.

Satz. *Sind in einem beschränkten Teil (Simplex oder konvexen Polyeder) des reellen euklidischen R^n endlichviele algebraische Varietäten oder algebraisch-begrenzte Teilbereiche von algebraischen Varietäten gegeben, so gibt es eine Unterteilung dieses Raumteils derart, daß alle gegebenen algebraischen Varietäten aus Seiten der Zellen der Unterteilung bestehen.*

Beweis. Induktion nach n. (Für $n = 0$ ist nichts zu beweisen.)

Die Gleichungen der gegebenen Varietäten sind, eventuell nach einer linearen Koordinatentransformation, so zu schreiben, daß sie regulär in x_n sind (d. h. daß die höchste Potenz von x_n einen konstanten Koeffizienten $\neq 0$ hat). Die Projektionen dieser Varietäten (soweit ihre Dimension n ist) und ihrer Randgebilde auf den Raum $x_n = 0$ sind algebraische Gebilde von derselben Art, und zu jedem Punkt der Projektion gehören endlichviele Originalpunkte, algebraische Funktionen der Projektion. Wir verzeichnen im Raum $x_n = 0$ auch alle Verzweigungsgebilde dieser algebraischen Funktionen (wo zwei Wurzeln x_n zusammenrücken), wobei diejenigen Teilgebilde, wo noch mehr Wurzeln zusammenrücken als sonst auf dem Verzweigungsgebilde, noch besonders verzeichnet werden. Auf die Gesamtheit der so erhaltenen algebraischen Gebilde des Raumes $x_n = 0$ wird die Induktionsvoraussetzung angewandt. Über jeder Zelle oder Seite der damit erhaltenen Unterteilung des betreffenden Teils des Raumes $x_n = 0$ liegt eine zylindrische Punktmenge des R^n, und auf jeder erzeugenden Geraden ($x_1 =$ konst., $\ldots$, $x_{n-1} =$ konst.) eines solchen Zylinders sind endlichviele zu den gegebenen algebraischen Varietäten gehörige x_n-Werte ausgezeichnet, welche die erzeugende Gerade also in endlichviele Strecken teilen. Ändert sich die erzeugende Gerade, so ändern sich diese Teilpunkte stetig, und es fallen nie (außer am Rande der Zelle) zwei Teilpunkte zusammen. Jede dieser Strecken durchläuft also eine Zelle des R^n. Damit ist die verlangte Zelleneinteilung des R^n konstruiert.

Durch eine baryzentrische Unterteilung erhält man, wenn man das wünscht, aus der Zelleneinteilung eine Triangulation oder Simplexeinteilung.

Mutatis mutandis läßt sich das alles auf analytische Varietäten übertragen. Wichtiger für uns ist aber die *Übertragung auf algebraische Varietäten im komplexen projektiven Raum.*

Es gibt, wenn man die komplexen Koordinaten $z_0, \ldots, z_n$ durch $\sum z_j \bar{z}_j = 1$ normiert, die folgende (bekannte) eineindeutige Abbildung des komplexen projektiven Raumes auf eine ganz im Endlichen liegende alge-

braische Varietät im reellen euklidischen Raum von n^2 Dimensionen:

$$z_j \bar{z}_k = x_{jk} + i y_{jk} \qquad (x_{jk} = x_{kj},\ y_{jk} = -y_{kj}).\ ^{29})$$

Bei dieser Abbildung gehen natürlich algebraische Varietäten von r komplexen Dimensionen in reelle algebraische Varietäten von $2r$ Dimensionen über. Der vorige Satz liefert daher die Möglichkeit einer Triangulation des komplexen projektiven Raumes, wobei beliebig vorgegebene algebraische Varietäten in ihm mit trianguliert werden können.

Anhang 2.

Die Topologie des reellen und des komplexen projektiven Raumes.

Zelleneinteilung. Die bequemste Zelleneinteilung des *reellen projektiven Raums* (homogene Koordinaten $z_0, \ldots, z_n$) erhält man so: Eine Zelle hat die Gleichungen

$$|z_1| \leqq |z_0|, \ldots, |z_n| \leqq |z_0|,$$

und die übrigen erhält man daraus durch Permutation der z_j. Verwandelt man einige der Ungleichungen in Gleichungen, so erhält man die Seiten der verschiedenen Dimensionen. Alle diese Zellen sind homöomorph Parallelotopen des euklidischen Raumes (man kann ja z. B. für alle Punkte der oben angeschriebenen Zelle die Größen $\frac{z_1}{z_0}, \ldots, \frac{z_n}{z_0}$ als inhomogene Koordinaten benutzen).

Im *komplexen projektiven Raum* hat man den obigen Ungleichungen noch andere von der Form $0 \leqq \arg\left(\frac{z_j}{z_0}\right) \leqq \pi$ für einige j, $\pi \leqq \arg\left(\frac{z_j}{z_0}\right) \leqq 2\pi$ für die übrigen j aus der Reihe $1, \ldots, n$ hinzuzufügen. Die Zellen werden topologische Produkte von Halbkreisen und Strecken. Bildet man eine baryzentrische Unterteilung, so sind alle Simplizes um einen Punkt so angeordnet wie Simplizes um einen Punkt im euklidischen Raum, wie man durch eine passende Abbildung leicht feststellt; der Raum ist mithin eine Mannigfaltigkeit im kombinatorischen Sinn[30]).

Die *Orientierbarkeit* des komplexen projektiven Raums wird genau so bewiesen wie die Orientierbarkeit aller algebraischen Varietäten im § 7:

[29]) Vgl. G. Mannoury, Surfaces-images, Nieuw Archief von Wiskunde (2) 4 (1898), p. 112.

[30]) Die etwas langwierige Verifikation dieser Behauptungen wird überflüssig gemacht durch die in Fußnote [20]) erwähnten neueren Untersuchungen. Denen zufolge genügt es ja, erstens nachzuweisen, daß es eine Triangulation überhaupt gibt, was in Anhang 1 schon geschah, und zweitens zu zeigen, daß der komplexe projektive Raum eine Mannigfaltigkeit im topologischen Sinn (vgl. § 7) ist. Letzteres ist klar, denn eine Umgebung eines jeden Punktes ist einem euklidischen Raum homöomorph: man kann für $z_0 \neq 0$ ja die Real- und Imaginärteile von $\frac{z_1}{z_0}, \ldots, \frac{z_n}{z_0}$ als Parameter benutzen.

Für einen Raumteil $|z_j| \leqq |z_0|$ $(j = 1, \ldots, n)$ kann man $\frac{z_1}{z_0}, \ldots, \frac{z_n}{z_0}$ als komplexe Parameter benutzen und den Raumteil entsprechend orientieren. Der Übergang zu einem angrenzenden Raumteil, wo etwa $\frac{z_0}{z_1}, \frac{z_2}{z_1}, \ldots, \frac{z_n}{z_1}$ als Parameter benutzt werden, vollzieht sich dann durch eine analytische Parametertransformation, und diese erhält die Orientierung.

Der reelle projektive Raum ist bekanntlich nur für die ungeraden Dimensionszahlen orientierbar.

Die Aufstellung einer *Homologiebasis* kann für den komplexen und den reellen Fall gemeinsam geschehen. Wir nehmen etwa den komplexen $P_n = P^{2n}$. Eine Basis für die Zyklen der Dimension $2n$ bildet natürlich der Raum P^{2n} selbst (im reellen Fall der Raum P^n, falls er orientierbar ist, also für ungerades n). Jeder Zykel C^k $(k < 2n)$ kann zunächst so deformiert werden, daß er den Punkt $(1, 0, \ldots, 0)$ nicht enthält. Deformiert man ihn jetzt weiter, indem man jeden Punkt $(x_0, \ldots, x_n)$ von C^k durch $(\lambda x_0, x_1, \ldots, x_n)$ ersetzt und die reelle Zahl λ von 1 nach 0 gehen läßt, so wird C^k schließlich ganz in den Unterraum $z = 0$ geschoben. Nimmt man nun als Induktionsvoraussetzung für den P_{n-1} an, daß in ihm eine Homologiebasis für alle Dimensionen durch die linearen Unterräume $L_j = L^{2j}$ $(j = 0, \ldots, n-1)$ gebildet wird, so folgt, daß unser Zykel Z^k für ungerades k homolog Null, für gerades $k = 2j$ homolog einem Vielfachen eines L_j im Unterraum $x_0 = 0$ ist. Damit ist die Induktionsvoraussetzung auch für den P_n bewiesen. Für P_0 ist sie trivial, mithin:

Im komplexen $P_n = P^{2n}$ bilden die linearen Unterräume $L_j = L^{2j}$ (inklusive P_n selbst) eine Homologiebasis für alle Dimensionen.

Daß die L_j sowie ihre Vielfache nicht homolog Null sind, sieht man daraus, daß die Schnittpunktszahl $L_j \cdot L_{n-j}$ den Wert 1 hat.

Im reellen projektiven P_n kommen die L_j als Zyklen nur für ungerades j in Betracht, da sie für gerades j ja nicht orientierbar sind; sie sind, doppelt gezählt, Rand eines L_{j+1}, mithin

$$2 L_j \sim 0.$$

Daß die L_j $(j = 2n + 1)$ selbst nicht homolog Null und sogar nicht homolog Null modulo 2 sind, sieht man daraus, daß die „Schnittpunktszahl mod 2“ von L_j und L_{n-j} (mod 2 ist auch L_{n-j} ein Zykel) den Wert 1 hat.

Der Grenzübergang $\lambda \to 0$, der hier zur Bestimmung der Homologiebasis führte, entspricht der „ausgearteten Transformation“ von W_2, § 7. Was aber algebraisch sehr kompliziert durchzuführen war, wird topologisch äußerst einfach.

(Eingegangen am 23. 2. 1929.)

Über den Aufbau des Nullideals in ganz abgeschlossenen Ringen mit Teilerkettensatz [1]).

Von

Wolfgang Krull in Erlangen.

Im folgenden wird gezeigt, daß in einem ganz abgeschlossenen Ring mit Teilerkettensatz die Struktur des Nullideals keineswegs willkürlich vorgeschrieben werden kann, sondern daß bei einem derartigen Ring — von einem trivialen Ausnahmefall abgesehen — unter den isolierten Primärkomponenten [2]) des Nullideals stets Primideale vorkommen müssen. Insbesondere ergibt sich, daß ein ganz abgeschlossener Ring mit Teilerkettensatz, in dem das Nullideal Durchschnitt seiner isolierten Primärkomponenten ist, stets dargestellt werden kann als direkte Summe von endlich vielen speziellen ganz abgeschlossenen Ringen, die entweder Integritätsbereiche

---

[1]) Die vorliegende Note stellt eine Ergänzung zweier kürzlich erschienenen Arbeiten von B. L. van der Waerden dar:

B. L. van der Waerden, Zur Produktzerlegung der Ideale in ganz abgeschlossenen Ringen, Math. Annalen **101** (1929), zitiert mit „W. I".

B. L. van der Waerden: Zur Idealtheorie der ganz abgeschlossenen Ringe, ebenda, zitiert mit „W. II".

Vgl. ferner: F. C. Schmidt, Primidealzerlegung für die Hauptideale eines Integritätsbereichs, Sitzungsber. d. Münchner Akademie 1929, wo ohne Beweis der folgende Satz angegeben ist:

In einem beliebigen Integritätsbereich J läßt sich dann und nur dann jedes Hauptideal eindeutig als Produkt von Primidealpotenzen darstellen, wenn 1. J ganz abgeschlossen ist, und wenn 2. in J jede mit einem Hauptideal beginnende Quotientenkette nur endlich viel verschiedene Glieder besitzt.

(Die zweite Bedingung ist wesentlich schwächer als der in der vorliegenden Note und in W. I u. W. II vorausgesetzte Noethersche Teilerkettensatz.)

[2]) Im folgenden benützen wir für die isolierten Komponenten eines Ideals die Definition, die van der Waerden in § 2 der Arbeit: Eine Verallgemeinerung des Bézontschen Theorems [Math. Annalen **99** (1928)], gegeben hat. Vgl. auch den Beginn von § 1 des Textes, wo die zum Verständnis unserer Untersuchungen notwendigen Tatsachen über isolierte Komponentenideale kurz zusammengestellt sind.

sind (d. h. außer 0 selbst keine Nullteiler enthalten) oder nur aus Nullteilern und Einheiten bestehen.

Auf Grund dieser Tatsache können die „Multiplikationsringe“, in denen aus der Teilbarkeit von $\mathfrak{a}$ durch $\mathfrak{b}$ stets die Existenz einer Produktdarstellung $\mathfrak{a} = \mathfrak{b} \cdot \mathfrak{c}$ folgt, idealtheoretisch erschöpfend charakterisiert werden. Schließlich wird noch durch ein Beispiel gezeigt, daß es ganz abgeschlossene Ringe mit Teilerkettensatz gibt, in denen das Nullideal nicht Durchschnitt seiner isolierten Primärkomponenten ist, für die also eine direkte Summenzerlegung der oben angegebenen Art nicht existiert.

§ 1.

Ein Satz über quasiumkehrbare Primideale.

Es sei $\mathfrak{R}$ ein — vorerst nicht notwendig ganz abgeschlossener — Ring mit Einheitselement ε, in dem der Teilerkettensatz gilt; $\mathfrak{o}$ bedeute das Einheits-, $\mathfrak{n}$ das Nullideal. $\mathfrak{K}$ sei der zu $\mathfrak{R}$ gehörige Quotientenring, also der (bis auf Isomorphie) eindeutig bestimmte kleinste Erweiterungsring von $\mathfrak{R}$, in dem jeder Nichtnullteiler von $\mathfrak{R}$ Einheit ist, d. h. ein hinsichtlich ε reziprokes Element besitzt. Die *isolierten Komponenten* eines Ideals $\mathfrak{m}$ definieren wir in üblicher Weise, also z. B. im Anschluß an van der Waerden: Ist S irgendein Elementesystem aus $\mathfrak{R}$, das gleichzeitig mit zwei Elementen stets auch deren Produkt enthält, so soll unter der durch S erzeugten isolierten Komponente $\mathfrak{m}_S$ von $\mathfrak{m}$ das Ideal verstanden werden, das alle und nur die Elemente aus $\mathfrak{R}$ enthält, deren Produkt mit einem jeweils geeignet gewählten Element aus S zu $\mathfrak{m}$ gehört. — Wegen der Gültigkeit des Teilerkettensatzes in $\mathfrak{R}$ kann man zu $\mathfrak{m}_S$ ein Element s aus S stets derart bestimmen, daß das Produkt von $\mathfrak{m}_S$ mit dem Hauptideal (s) Vielfaches von $\mathfrak{m}$ wird. — Ist $\mathfrak{m}_S$ Primärideal, so nennen wir $\mathfrak{m}_S$ eine *isolierte Primärkomponente* von $\mathfrak{m}$. Das zu einer isolierten Primärkomponente gehörige Primideal ist *höchstes Primideal* von $\mathfrak{m}$, d. h. es ist zwar selbst Teiler von $\mathfrak{m}$, besitzt aber kein echtes Primidealvielfaches mit der gleichen Eigenschaft. Umgekehrt gehört zu einem höchsten Primideal von $\mathfrak{m}$ stets eine bestimmte isolierte Primärkomponente.

Ein Ideal $\mathfrak{a}$ aus $\mathfrak{R}$, das mindestens einen Nichtnullteiler enthält, soll *regulär* heißen. Wegen der Gültigkeit des Teilerkettensatzes ist $\mathfrak{a}$ dann und nur dann regulär, wenn $\mathfrak{a}$ zum Nullideal prim ist, wenn also kein von $\mathfrak{n}$ verschiedenes Ideal $\mathfrak{m}$ existiert, das der Gleichung $\mathfrak{a} \cdot \mathfrak{m} = \mathfrak{n}$ genügt[3]). Wir werden von dieser Tatsache öfters Gebrauch machen.

---

[3]) Vgl. § 1 S. 5 der Arbeit: Krull, Primidealketten in allgemeinen Ringbereichen, Sitzungsber. d. Heidelberger Akademie, Math.-Naturw. Klasse, 1928, 3. Abhandl., zitiert mit K. I.

Ist $\mathfrak{p} \neq \mathfrak{o}$ ein beliebiges reguläres Primideal aus $\mathfrak{R}$, so verstehen wir unter dem gebrochenen Ideal $\mathfrak{o} : \mathfrak{p}$ in üblicher Weise die Gesamtheit derjenigen Elemente aus dem Quotientenring $\mathfrak{K}$, deren Produkte mit den Elementen von $\mathfrak{p}$ sämtlich zu $\mathfrak{R}$ gehören. Ist $\mathfrak{p}\cdot(\mathfrak{o}:\mathfrak{p}) = \mathfrak{o}$, so heißt $\mathfrak{p}$ *umkehrbar*, stellt allgemeiner $\mathfrak{p}\cdot(\mathfrak{o}:\mathfrak{p})$ irgendeinen echten Teiler von $\mathfrak{p}$ dar, so bezeichnen wir $\mathfrak{p}$ als *quasiumkehrbar*. Es gilt nun

Satz 1. *Ist das reguläre Primideal $\mathfrak{p}$ aus $\mathfrak{R}$ quasiumkehrbar, so gibt es in $\mathfrak{R}$ nur ein einziges echtes Primidealvielfaches $\mathfrak{p}_n$ von $\mathfrak{p}$ und es stellt dabei $\mathfrak{p}_n$ eine isolierte Primärkomponente des Nullideals dar.*

Das System $S_\mathfrak{p}$ aller durch $\mathfrak{p}$ unteilbarer Elemente aus $\mathfrak{R}$ enthält gleichzeitig mit a und b stets auch $a\cdot b$, $S_\mathfrak{p}$ erzeugt daher eine isolierte Komponente $\mathfrak{n}_\mathfrak{p}$ des Nullideals. Da nach einer oben gemachten Bemerkung zu $\mathfrak{n}_\mathfrak{p}$ ein durch $\mathfrak{p}$ unteilbares Hauptideal (s) existiert, das der Gleichung $\mathfrak{n}_\mathfrak{p}\cdot(s) = \mathfrak{n}$ genügt, ist $\mathfrak{n}_\mathfrak{p}$ durch jedes Primidealvielfache von $\mathfrak{p}$ teilbar. Andrerseits folgt aus früher von mir angestellten Untersuchungen: Bedeutet $\mathfrak{p}^{(\nu)}$ die durch $S_\mathfrak{p}$ erzeugte zu $\mathfrak{p}$ gehörige isolierte Primärkomponente von $\mathfrak{p}^\nu$, so ist der Durchschnitt der sämtlichen Ideale $\mathfrak{p}^{(\nu)}$ gerade gleich $\mathfrak{n}_\mathfrak{p}$.[4]) Zum Beweise von Satz 1 brauchen wir daher nur zu zeigen:

Jedes echte Primidealvielfache $\mathfrak{p}_n$ von $\mathfrak{p}$ ist durch alle $\mathfrak{p}^{(\nu)}$ teilbar.

In der Tat, zunächst ist nach Voraussetzung $\mathfrak{p}_n$ Vielfaches von $\mathfrak{p} = \mathfrak{p}^{(1)}$. Wir nehmen nun an, es sei $\mathfrak{p}_n$ durch $\mathfrak{p}^{(\nu)}$ teilbar und erschließen daraus die Teilbarkeit von $\mathfrak{p}_n$ durch $\mathfrak{p}^{(\nu+1)}$ auf folgende Weise[5]): $\mathfrak{p}_n\cdot(\mathfrak{o}:\mathfrak{p})$ ist ein (ganzes) Ideal aus $\mathfrak{R}$, und es ist $(\mathfrak{p}_n\cdot(\mathfrak{o}:\mathfrak{p}))\cdot\mathfrak{p} = \mathfrak{p}_n\cdot((\mathfrak{o}:\mathfrak{p})\cdot\mathfrak{p})$ durch $\mathfrak{p}_n$ teilbar. Da nach Voraussetzung $\mathfrak{p}_n$ echtes Vielfaches von $\mathfrak{p}$ ist, ergibt sich daraus die Teilbarkeit von $\mathfrak{p}_n\cdot(\mathfrak{o}:\mathfrak{p})$ durch $\mathfrak{p}_n$, es stellt daher $\mathfrak{p}_n\cdot(\mathfrak{o}:\mathfrak{p})\cdot\mathfrak{p}$ ein Vielfaches von $\mathfrak{p}_n\cdot\mathfrak{p}$ dar. Beachten wir nun, daß $\mathfrak{p}_n\cdot\mathfrak{p}$ durch das zu $\mathfrak{p}$ gehörige Primärideal $\mathfrak{p}^{(\nu+1)}$ teilbar, $(\mathfrak{o}:\mathfrak{p})\cdot\mathfrak{p}$ dagegen durch $\mathfrak{p}$ unteilbar ist, so folgt aus dem zuletzt gewonnenen Ergebnis die Teilbarkeit von $\mathfrak{p}_n$ durch $\mathfrak{p}^{(\nu+1)}$. Damit ist gezeigt, daß $\mathfrak{p}_n$ gemeinsames Vielfaches von sämtlichen $\mathfrak{p}^{(\nu)}$ sein muß, der Beweis von Satz 1 ist beendet.

---

[4]) K. I. § 2. Dort habe ich nämlich einerseits beim Beweise von Satz 3 gezeigt: „Der Durchschnitt $\mathfrak{m}$ aller $\mathfrak{p}^{(\nu)}$ genügt der Gleichung $\mathfrak{m}\cdot\mathfrak{p} = \mathfrak{m}$" und andrerseits beim Beweise von Satz 2 das Ergebnis gewonnen: „Ist $\mathfrak{m}\cdot\mathfrak{p} = \mathfrak{m}$, so gibt es sicher ein durch $\mathfrak{p}$ unteilbares Hauptideal (s), das der Gleichung $\mathfrak{m}\cdot(s) = \mathfrak{n}$ genügt." Daraus ergibt sich die Teilbarkeit von $\mathfrak{m}$ durch $\mathfrak{n}_\mathfrak{p}$, und da umgekehrt $\mathfrak{n}_\mathfrak{p}$ ebenso wie durch jedes Primidealvielfache von $\mathfrak{p}$ auch durch jedes zu $\mathfrak{p}$ gehörige Primärideal teilbar ist, muß $\mathfrak{n}_\mathfrak{p} = \mathfrak{m}$ sein. Daß übrigens die $\mathfrak{p}^{(\nu)}$ wirklich zu $\mathfrak{p}$ gehörige Primärideale darstellen, folgt aus dem Umstand, daß einerseits jeder Primidealteiler von $\mathfrak{p}^{(\nu)}$ Teiler von $\mathfrak{p}$ sein muß, und daß andrerseits wegen der Art der Erzeugung von $\mathfrak{p}^{(\nu)}$ aus $a\cdot b \equiv 0(\mathfrak{p}^{(\nu)})$; $b \not\equiv 0(\mathfrak{p})$ stets $a \equiv 0(\mathfrak{p}^{(\nu)})$ folgt.

[5]) Auf die folgenden Schlüsse, die den Beweis wesentlich abkürzen, machte mich Herr van der Waerden aufmerksam.

§ 2.

Anwendung von Satz 1 auf ganz abgeschlossene Ringe.

Wir wenden jetzt Satz 1 auf den Fall an, daß $\mathfrak{R}$ im Quotientenring $\mathfrak{K}$ ganz abgeschlossen ist. Existiert hier in $\mathfrak{R}$ ein Element a, das weder Einheit noch Nullteiler ist, so gibt es in $\mathfrak{R}$ nach v. d. Waerden [6]) mindestens ein quasiumkehrbares Primideal, es ist nämlich z. B. jedes zum Hauptideal (a) gehörige Primideal quasiumkehrbar. Die Anwendung von Satz 1 liefert daher:

Satz 2. *In einem ganz abgeschlossenen Ring, in dem der Teilerkettensatz gilt, ist entweder jeder Nichtnullteiler Einheit, oder es ist mindestens eine isolierte Primärkomponente des Nullideals Primideal.*

Spezialisieren wir Satz 2 auf den Fall eines „primären" Ringes $\mathfrak{R}$, in dem das Nullideal Primärideal ist, also nur ein einziges zugehöriges Primideal besitzt, so erhalten wir:

Satz 3. *Ein primärer ganz abgeschlossener Ring $\mathfrak{R}$, in dem der Teilerkettensatz gilt, ist entweder ein Integritätsbereich oder es gibt in $\mathfrak{R}$ nur Nullteiler und Einheiten.*

Auf Satz 3 läßt sich schließlich noch der Fall zurückführen, daß in $\mathfrak{R}$ das Nullideal den Durchschnitt von endlich vielen, gegenseitig primen Primäridealen $\mathfrak{q}_1, \mathfrak{q}_2, \ldots, \mathfrak{q}_s$, also den Durchschnitt seiner isolierten Primärkomponenten darstellt. Verstehen wir hier unter $\mathfrak{q}_i^*$ das Ideal $\mathfrak{q}_i \cdot \mathfrak{K}$ aus $\mathfrak{K}$, so sind die $\mathfrak{q}_i^*$ paarweise teilerfremd, weil ja unter unserer Voraussetzung der gr. g. T. von $\mathfrak{q}_i$ und $\mathfrak{q}_k$ in $\mathfrak{R}$ für $i \neq k$ stets einen Nichtnullteiler enthält, der in $\mathfrak{K}$ Einheit wird. Da außerdem in $\mathfrak{K}$ der Durchschnitt der Ideale $\mathfrak{q}_i^*$ gleich dem Nullideal ist, gilt nach bekannten Sätzen [7]) für $\mathfrak{K}$ eine direkte Summenzerlegung $\mathfrak{K} = \mathfrak{K}_1 + \mathfrak{K}_2 + \ldots + \mathfrak{K}_s$, wobei der Ring $\mathfrak{K}_i$ aus den Elementen des Ideals $\mathfrak{q}^{(i)*} = \prod\limits_{k \neq i} \mathfrak{q}_k^*$ besteht und zu $\mathfrak{K} \,|\, \mathfrak{q}_i^*$ isomorph ist. Bedeutet ε_i das Einheitselement aus $\mathfrak{K}_i$, so ist ε_i wegen $\varepsilon_i^2 - \varepsilon_i = 0$ von $\mathfrak{R}$ ganz abhängig und mithin schon im Durchschnitt von $\mathfrak{R}$ und $\mathfrak{K}_i = \mathfrak{q}_i^*$, d. h. im Durchschnitt der Ideale $\mathfrak{q}_1, \ldots, \mathfrak{q}_{i-1}, \mathfrak{q}_{i+1}, \ldots, \mathfrak{q}_s$ enthalten. Daraus folgt angesichts der Gleichung $\varepsilon = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_s$, daß bereits in $\mathfrak{R}$ die Ideale $\mathfrak{q}_1, \mathfrak{q}_2, \ldots, \mathfrak{q}_s$ paarweise teilerfremd sind.

Der direkten Summendarstellung $\mathfrak{K} = \mathfrak{K}_1 + \mathfrak{K}_2 + \ldots + \mathfrak{K}_s$ entspricht daher eine direkte Summendarstellung $\mathfrak{R} = \mathfrak{R}_1 + \mathfrak{R}_2 + \ldots + \mathfrak{R}_s$, $\mathfrak{R}_i = \prod\limits_{k \neq i} \mathfrak{q}_k = \mathfrak{q}^{(i)}$. Dabei stellt offenbar $\mathfrak{K}_i$ gerade den zu $\mathfrak{R}_i$ gehörigen Quotientenring dar,

---

[6]) W. II: Beweis des Satzes v. S. 309.

[7]) Was die Theorie der eindeutigen additiven Zerlegung angeht, so vgl. z. B. die Zusammenstellung in § 3 der Arbeit: E. Noether, Der Diskriminantensatz für die Ordnungen eines algebraischen Zahl- oder Funktionenkörpers, Journ. f. Math. 157 (1926).

und es ist $\mathfrak{R}_i$ gleich dem Durchschnitt von $\mathfrak{K}_i$ und $\mathfrak{R}$. Hängt nun α_i aus $\mathfrak{K}_i$ von $\mathfrak{R}_i$ ganz ab, so hängt es erst recht von $\mathfrak{R}$ ganz ab, es gehört also zu $\mathfrak{R}$ und damit auch zu $\mathfrak{R}_i$, d. h.: Die Ringe $\mathfrak{R}_i$ sind ganz abgeschlossen. Da jeder $\mathfrak{R}_i$ wegen seiner Isomorphie zu $\mathfrak{R}|\mathfrak{q}_i$ primär ist, können wir Satz 3 anwenden und wir erhalten:

Satz 4. *Genügt der ganz abgeschlossene Ring $\mathfrak{R}$ dem Teilerkettensatz, und ist in $\mathfrak{R}$ das Nullideal Durchschnitt seiner isolierten Primärkomponenten, so ist $\mathfrak{R}$ die direkte Summe von endlich vielen primären ganz abgeschlossenen Ringen, die entweder Integritätsbereiche sind oder nur Nullteiler und Einheiten enthalten.*

Wir zeigen jetzt durch ein Beispiel, daß nicht alle ganz abgeschlossenen Ringe mit Teilerkettensatz unter den in Satz 4 erledigten Spezialfall gehören, d. h. wir beweisen: *Es gibt ganz abgeschlossene Ringe, die dem Teilerkettensatz genügen und in denen das Nullideal von dem Durchschnitt seiner isolierten Primärkomponenten verschieden ist.*

Es sei K ein beliebiger Körper, $\mathfrak{S} = K[x, y]$ sei der Ring aller Polynome in x und y mit Koeffizienten aus K, $\mathfrak{R}$ sei das Restklassensystem von $\mathfrak{S}$ nach dem Ideal (x^2, xy). Dann genügt $\mathfrak{R}$ dem Teilerkettensatz, und es ist das Nullideal von $\mathfrak{R}$ von dem Durchschnitt seiner isolierten Primärkomponenten verschieden, weil es dieselbe Struktur besitzt wie das Ideal (x^2, xy) aus $\mathfrak{S}$, das (x) als einzige isolierte Primärkomponente hat.

Schließlich ist $\mathfrak{R}$ ganz abgeschlossen.

In der Tat, es seien $\bar{p}, \bar{q}, \bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$ beliebige Elemente aus $\mathfrak{R}$, $\bar{q}$ insbesondere Nichtnullteiler, und es hänge $\bar{a} = \frac{\bar{p}}{\bar{q}}$ vermöge der Gleichung $\bar{a}^n + \bar{a}_1 \bar{a}^{n-1} + \ldots + \bar{a}_n = 0$, d. h. vermöge $\bar{p}^n + \bar{a}_1 \bar{p}^{n-1} \bar{q} + \ldots + \bar{a}_n \bar{q}^n = 0$ von $\mathfrak{R}$ ganz ab. Verstehen wir dann unter $p(x, y) = a \cdot x + \varphi(y)$ bzw. $q(x, y) = b \cdot x + \psi(y)$ einen Vertreter der Restklasse $\bar{p}$ bzw. $\bar{q}$, so folgt in $\mathfrak{S}$ aus der für $\bar{p}$ und $\bar{q}$ gültigen Gleichung eine Kongruenz: $\varphi^n(y) + a_1(y)\, \varphi^{n-1}(y) \cdot \psi(y) + \ldots + a_n(y)\, \psi^n(y) \equiv 0\ ((x))$, die uns zeigt, daß $\varphi(y)$ durch $\psi(y)$ in $\mathfrak{S}$ teilbar sein muß, $\varphi(y) = \psi(y) \cdot \chi(y)$. Da ferner $\bar{q}$ in $\mathfrak{R}$ Nichtnullteiler ist, muß $q(x, y) = b \cdot x + \psi(y)$ durch das Ideal (x, y) unteilbar, d. h. es muß $\psi(y)$ modulo y einem von 0 verschiedenen Element aus K kongruent sein. Es gibt daher in K sicher eine Lösung d der Kongruenz $d \cdot \psi(y) + b \cdot \chi(y) \equiv a\ ((y))$. Setzt man nun $r(x, y) = d \cdot x + \chi(y)$, so wird $q(x, y) \cdot r(x, y) \equiv p(x, y)\ ((x^2, xy))$, d. h. es genügt die durch $r(x, y)$ definierte Restklasse $\bar{r}$ aus $\mathfrak{R}$ der Gleichung $\bar{p} = \bar{q} \cdot \bar{r}$, $\frac{\bar{p}}{\bar{q}} = \bar{r}$ gehört zu $\mathfrak{R}$. Wir haben also aus der Tatsache, daß $\bar{a}$ von $\mathfrak{R}$ ganz abhängt, die Zugehörigkeit von $\bar{a}$ zu $\mathfrak{R}$ erschlossen.

Da $\bar{p}$ ein ganz beliebiges Element, $\bar{q}$ einen ganz beliebigen Nichtnullteiler aus $\mathfrak{R}$ bedeutete, ist damit die ganze Abgeschlossenheit von $\mathfrak{R}$ bewiesen, und es ist gezeigt, daß $\mathfrak{R}$ wirklich ein Beispiel der von uns gewünschten Art darstellt.

§ 3.

Über Multiplikationsringe[8]).

Ein Ring $\mathfrak{R}$, in dem der Teilerkettensatz gilt, soll *Multiplikationsring* heißen, wenn in $\mathfrak{R}$ aus der Teilbarkeit von $\mathfrak{a}$ durch $\mathfrak{b}$ stets die Gültigkeit einer Gleichung $\mathfrak{a} = \mathfrak{b} \cdot \mathfrak{c}$ folgt. Für Multiplikationsringe gelten zunächst folgende Hilfssätze:

Hilfssatz 1. *In einem Multiplikationsring läßt sich jedes reguläre Ideal als Produkt von Potenzen teilerfremder Primideale darstellen*[9]).

Hilfssatz 2. *Besitzt in einem Multiplikationsring das nichtreguläre Ideal* $\mathfrak{a}$ *einen gleichfalls nichtregulären echten Teiler* $\mathfrak{b}$, *so läßt sich* $\mathfrak{a}$ *als Produkt echter Teiler darstellen*[10]).

Den Hilfssatz 1 beweist man ohne Schwierigkeit auf Grund der Definition des Multiplikationsringes und der Tatsache, daß in einem Ring mit Teilerkettensatz aus $\mathfrak{a} = \mathfrak{a} \cdot \mathfrak{b}$, $\mathfrak{a}$ regulär, stets $\mathfrak{b} = \mathfrak{o}$ folgt. Um Hilfssatz 2 zu beweisen, schließen wir so: Aus der Teilbarkeit von $\mathfrak{a}$ durch $\mathfrak{b}$ folgt die Gültigkeit einer Gleichung $\mathfrak{a} = \mathfrak{b} \cdot \mathfrak{c}$; bedeutet ferner $\mathfrak{d}$ irgendein wegen der Nichtregularität von $\mathfrak{b}$ sicher existierendes, von $\mathfrak{n}$ verschiedenes Ideal, das der Gleichung $\mathfrak{b} \cdot \mathfrak{d} = \mathfrak{n}$ genügt, so dürfen wir $\mathfrak{c}$ als Teiler von $\mathfrak{d}$ annehmen. Unter dieser Voraussetzung aber wird $\mathfrak{c}$ echter Teiler von $\mathfrak{a}$; wäre nämlich $\mathfrak{c} = \mathfrak{a}$, so wäre $\mathfrak{a}$ Teiler von $\mathfrak{d}$, und daraus folgte $\mathfrak{d} = \mathfrak{e} \cdot \mathfrak{a} = \mathfrak{e} \cdot \mathfrak{b} \cdot \mathfrak{a} = \mathfrak{b} \cdot \mathfrak{d} = \mathfrak{n}$.

Aus Hilfssatz 1 ergibt sich nach einem van der Waerdenschen Satz[11]) die ganze Abgeschlossenheit der Multiplikationsringe. Hilfssatz 2 zeigt, daß in einem Multiplikationsring zwei nichtreguläre Primideale stets gegenseitig prim sind. Wir können daher Satz 4 anwenden, und wir erhalten:

---

[8]) Die in § 3 über Multiplikationsringe abgeleiteten Sätze habe ich schon früher bewiesen, und zwar in den beiden Noten:

Die verschiedenen Arten der Hauptidealringe, Sitzungsber. d. Heidelberger Akademie, Math.-naturw. Klasse, 1924, 6. Abhandl.; zitiert mit K. II.

Beiträge zur Algebra 3: Über Multiplikationsringe, Sitzungsber. d. Heidelberger Akademie, Math.-naturw. Klasse, 1925, 5. Abhandl.; zitiert mit K. III.

Im Text handelt es sich vor allem um die Einordnung der Multiplikationsringe in die allgemeine Theorie der ganz abgeschlossenen Ringe.

[9]) Vgl. K. II § 2, K. III S. 15.

[10]) Vgl. K. II § 3, Satz 3.

[11]) Vgl. W. II S. 310.

Satz 5. *Ein beliebiger Ring $\mathfrak{R}$ ist dann und nur dann Multiplikationsring, wenn er sich darstellen läßt als direkte Summe von endlich vielen speziellen Multiplikationsringen, die entweder Integritätsbereiche sind, oder primär sind und nur Nullteiler und Einheiten enthalten.*

(Man beachte, daß offenbar jeder direkte Summand eines Multiplikationsringes selbst wieder Multiplikationsring ist!) — Ein Integritätsbereich mit Multiplikationsringeigenschaft besitzt offenbar, wie aus Hilfssatz 1 zu ersehen, vom Standpunkt der Idealtheorie aus dieselbe Struktur wie der Integritätsbereich aller ganzen Zahlen eines endlichen algebraischen Zahlkörpers. Ein primärer Multiplikationsring dagegen, in dem nur Nullteiler und Einheiten auftreten, ist ein *„spezieller zerlegbarer Ring"*, d. h. er besitzt die idealtheoretische Struktur des Restklassensystems nach einer Primzahlpotenz [12]). Bedeutet nämlich $\mathfrak{p} \neq \mathfrak{n}$ das nichtreguläre Primideal eines derartigen Multiplikationsringes, π ein durch $\mathfrak{p}$, aber nicht durch $\mathfrak{p}^2$ teilbares Element, so gilt eine Gleichung $\pi^{\nu} = 0$, und es muß $\mathfrak{p} = (\pi)$ sein, weil das Ideal (π) das Produkt von $\mathfrak{p}$ mit einem sicher durch $\mathfrak{p}$ unteilbaren Faktor darstellen muß. Man erkennt nun ohne weiteres, daß jedes Ringelement auf die Form $\alpha = \pi^{s} \cdot \delta$ gebracht werden kann, wobei δ Einheit und der Exponent s für $\alpha \neq 0$ eindeutig bestimmt ist. Damit ist die Behauptung über die Struktur eines primären Multiplikationsringes bewiesen, wir sehen somit, daß durch Satz 5 das Studium eines beliebigen Multiplikationsringes auf die Untersuchung von wohlbekannten Ringtypen zurückgeführt wird.

---

[12]) Vgl. K. II S. 15, K. III S. 16 f.

(Eingegangen am 30. 1. 1929.)

Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren.

Von

J. v. Neumann in Berlin.

Einleitung.

1. Die vorliegende Arbeit zerfällt in zwei, im wesentlichen unabhängige, Teile. Der erste (§§ I—III) ist der Untersuchung der linearen und beschränkten Operatoren (d. h. Matrizen) des Hilbertschen Raumes $\mathfrak{H}$ gewidmet, indem die algebraischen Eigenschaften des von ihnen gebildeten (nichtkommutativen) Ringes $\mathfrak{B}$ betrachtet werden. Den Gegenstand des zweiten Teiles hingegen bilden diejenigen, nicht notwendig überall (in $\mathfrak{H}$) sinnvollen und beschränkten Operatoren, die die sogenannte Hilbertsche Spektraldarstellung mit komplexen Eigenwerten zulassen (vgl. die ausführlichere Explizierung dieser Begriffe im § 4 der Einleitung). Dies sind die als „normal" zu bezeichnenden Operatoren, die bisher nur im Beschränkten betrachtet wurden[1]), und für die wir eine neue allgemeinere Definition geben werden (vgl. am vorhin angeführten Orte).

Ehe wir diese Dinge genauer auseinandersetzen, sei an die Definition des (komplexen) Hilbertschen Raumes $\mathfrak{H}$ erinnert. Man kann ihn als die Menge aller Folgen komplexer Zahlen $\{x_1, x_2, \ldots\}$ mit endlichem $\sum_{n=1}^{\infty} |x_n|^2$ realisiert denken[2]); wir bezeichnen seine Elemente mit $f, g, \ldots, \varphi, \psi, \ldots$.

---

[1]) Bis vor kurzem nur im Vollstetigen, vgl. den Encyklopädie-Artikel von Hellinger und Toeplitz, Encykl. d. math. Wiss. II. C. 13, Seite 1562. Vgl. auch Anm. [25]).

[2]) Nach dem bekannten Satze von Fischer und F. Riesz ebensogut auch als Menge aller Funktionen $f(x)$ des Intervalles $a < x < b$ mit endlichem $\int_a^b |f(x)|^2 dx$ ($a < b$, jedes kann auch unendlich sein); oder aller Funktionen $f(P)$, wobei P die Einheitskugel-Oberfläche durchläuft, mit endlichem $\iint |f(P)|^2 d\sigma$ ($\iint \ldots d\sigma$ das Integral über die Einheitskugel-Oberfläche), usw. Vgl. z. B. die in Anm. [7]) zitierte Arbeit insbesondere Kap. I und Anhang I sowie Einleitung, zum Satz von Fischer und F. Riesz etwa F. Riesz, Göttinger Nachr. 1907, S. 210—273.

Mit diesen kann man wie mit Vektoren rechnen, so daß der Sinn von Bildungen wie $f \pm g$, af ohne weiteres einleuchtet[3]). Ferner gibt es ein „inneres Produkt" (f, g) wie bei Vektoren[4]), welches das Definieren des „absoluten Wertes" im Hilbertschen Raume ermöglicht: $|f| = \sqrt{(f, f)}$, und das Definieren der Entfernung $|f - g|$.[5]) Hierdurch wird der Hilbertsche Raum $\mathfrak{H}$ zu einem topologischen (ja metrischen) linearen Raume[6]), in welchem man geometrisch-topologische Redeweisen, wie lineare Mannigfaltigkeit, Häufungspunkt, Stetigkeit usw., ohne weiteres verwenden darf. Zur konsequenten Kennzeichnung von $\mathfrak{H}$ auf dieser Grundlage vergleiche man etwa mit einer früheren Arbeit des Verf.[7]), deren Bezeichnungen und Begriffsbildungen hier auch sonst verwendet werden sollen. Vgl. hierzu besonders die in Anm. [2]) genannten Teile derselben; immerhin werden wir die zur Anwendung kommenden Abschnitte von E. stets genau zitieren.

Die für unsere Zwecke wichtigsten Definitionen aus E. sind (Abkürzungen nach Anm. [7])): Ein O. ist eine Funktion, deren Wertevorrat und Definitionsbereich Teilmengen von $\mathfrak{H}$ sind — wir bezeichnen die O. mit $A, B, \ldots, R, S, \ldots$, den Wert von A an der Stelle f mit Af. (Af braucht also keineswegs überall in $\mathfrak{H}$ sinnvoll zu sein.) Er ist lin., wenn mit Af, Ag auch $A(af)$, $A(f+g)$ sinnvoll sind (d. h. der Definitionsbereich

---

[3]) Es wird definiert: $\{x_1, x_2, \ldots\} \pm \{y_1, y_2, \ldots\} = \{x_1 \pm y_1, x_2 \pm y_2, \ldots\}$, $a\{x_1, x_2, \ldots\} = \{ax_1, ax_2, \ldots\}$, wenn $\mathfrak{H}$ als „Folgenraum" der $\{x_1, x_2, \ldots\}$ (mit endlichem $\sum_{n=1}^{\infty} |x_n|^2$) interpretiert ist. Im „Funktionenraume" der $f(x)$ $\left(\int_a^b |f(x)|^2 dx \text{ endlich}\right)$ ist entsprechend $(f \pm g)(x) = f(x) \pm g(x)$, $(af)(x) = af(x)$; analog für andere Beispiele. Vgl. das Zitat von Anm. [2]).

[4]) Es wird definiert: $(\{x_1, x_2, \ldots\}, \{y_1, y_2, \ldots\}) = \sum_{n=1}^{\infty} x_n \bar{y}_n$ (die Reihe konvergiert absolut) bzw. $(f(x), g(x)) = \int_a^b f(x) \bar{g}(x)\, dx$, usw. Vgl. wie oben.

[5]) Somit ist $|\{x_1, x_2, \ldots\}| = \sqrt{\sum_{n=1}^{\infty} |x_n|^2}$ bzw. $|f(x)| = \sqrt{\int_a^b |f(x)|^2 dx}$ (leider ist hier die Bezeichnung mißverständlich, auf der linken Seite steht der Absolutwert der Funktion $f(x)$, rechts diejenigen ihrer Werte — bei unseren Überlegungen in $\mathfrak{H}$ wird aber so etwas nie vorkommen), usw. Vgl. wie oben.

[6]) Vgl. Hausdorff, Umgebungsaxiome und topologisch-lineare Räume: Grundzüge der Mengenlehre, I. Auflage, Leipzig und Berlin 1914.

[7]) „Zur allgemeinen Eigenwerttheorie Hermitescher Funktionaloperatoren", Math. Annalen **102**, 1 (1929), S. 49–131. Diese im folgenden mehrfach zu zitierende Arbeit soll stets mit E. bezeichnet werden. Von Wichtigkeit sind einige Abkürzungen aus E., die wir verwenden werden; sie seien hier zusammengestellt: Abgeschlossen = abg., beschränkt = beschr., Fortsetzung = Forts., Hermitesch = H., hypermaximal = hypermax., linear = lin., lineare Mannigfaltigkeit = lin. M., maximal = max., normiert = norm., Operator = O., orthogonal = orth., Projektion = P., vollständig = vollst., Zerlegung der Einheit = Z. d. E. Ferner werden wir den Begriff „vertauschbar" (vgl. Anm. [10])) mit vert. abkürzen.

eine lin. M. ist[8])), und zwar gleich aAf bzw. $Af+Ag$. Er heißt abg., wenn er die folgende Eigenschaft besitzt: Wenn alle Af_n $(n=1,2,\ldots)$ sinnvoll sind, und $f_n \to f$, $Af_n \to f^*$ (für $n \to \infty$) gilt, so ist Af sinnvoll und gleich f^* [9]).

Stetig ist ein O., wenn er es im Sinne der üblichen Definition in seinem ganzen Definitionsbereiche ist: Wenn Af_0 Sinn hat, so existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$, so daß aus $|f-f_0| < \delta$ $|Af - Af_0| < \varepsilon$ folgt (wenn Af Sinn hat). Man sieht sofort: Für einen abg. stetigen O. ist der Definitionsbereich notwendig eine abg. Menge — ist der O. auch lin., so ist der Definitionsbereich also eine abg. lin. M. [10]).

In der ersten Hälfte dieser Arbeit werden wir uns, wie schon angekündigt, mit überall sinnvollen, lin., stetigen O. beschäftigen, die (in Anlehnung an Hilbert) beschr. heißen sollen. (Vgl., auch für das Folgende, den Anhang I von E.) Für diese überall sinnvollen O. sind die einfachsten Rechenspezies so zu definieren: Wenn A, B beschr. sind, so sind es aA, $A \pm B$, AB auch, und zwar gilt

$$(aA)f = a \cdot Af, \quad (A \pm B)f = Af \pm Bf, \quad (AB)f = A(Bf).$$

Es gelten die Regeln der Algebra, mit zwei Ausnahmen: Es gibt Nullteiler, und die Multiplikation ist nicht kommutativ[11]). Die Menge $\mathfrak{B}$ aller beschr. O. ist also ein Ring — mit Nullteilern und nicht kommutativ. Eine weitere wichtige Operation ist $*$, die so definiert ist: Zu jedem beschr. A gibt es genau ein beschr. A^* mit

$$(Af, g) = (f, A^*g), \quad (f, Ag) = (A^*f, g).$$ [12])

---

[8]) D. h. mit f, g auch af, $f+g$ enthält. Eine abg. lin. M. muß auch noch abg. sein, d. h. alle ihre Häufungspunkte enthalten.

[9]) Diese stetigkeitsähnliche Eigenschaft ist in kompakten Räumen der Stetigkeit gleichwertig, aber in $\mathfrak{H}$ viel schwächer; vgl. E. Anm. [29]).

[10]) Für lin. O. kann die Stetigkeit auf mehrere, gleichwertige Formulierungen zurückgeführt werden, vgl. E. Satz 12. Eine von diesen ist: Es gebe ein festes c, so daß aus $|f| \leq 1$ $|Af| \leq c$ folgt.

[11]) 0, 1 sind durch $0f = 0$, $1f = f$ zu definieren.

[12]) Jede der beiden Relationen folgt, durch Vertauschen von f, g und Nehmen der Komplex-Konjugierten, aus der anderen. — Der Operation $*$ entspricht bei Matrizen das Nehmen der Komplex-Konjugiert-Transponierten. D. h. wenn $\mathfrak{H}$ als „Folgenraum" der $\{x_1, x_2, \ldots\}$ $\left(\sum_{n=1}^{\infty} |x_n|^2 \text{ endlich}\right)$ realisiert wird und A die Matrix $\{a_{\mu\nu}\}$ hat (also wenn stets

$$A\{x_1, x_2, \ldots\} = \{y_1, y_2, \ldots\},$$
$$y_\nu = \sum_{\mu=1}^{\infty} a_{\mu\nu} x_\mu$$

gilt — analog wie im Anhang II von E.), so hat A^* die Matrix $\{\overline{a_{\nu\mu}}\}$. Vgl. E. Anhang II.

Man verifiziert leicht die Regeln:

$$0^* = 0,\quad 1^* = 1,\quad (aA)^* = \bar{a}A^*,\quad (A \pm B)^* = A^* \pm B^*,\quad (AB)^* = B^* A^*.$$

Bei den unbeschr. (genauer: nicht notwendig beschr.) O., die den Gegenstand der zweiten Hälfte dieser Arbeit bilden, müssen wir etwas anders vorgehen. Für solche O. R, S ist zunächst bei der Definition von aR, $R \pm S$, RS zu beachten, daß die Definitionsbereiche dieser O. dadurch eingeengt werden, daß Rf, Sf nicht überall sinnvoll sein müssen. Damit $(R+S)f$ Sinn habe, müssen ja Rf, Sf sinnvoll sein, ebenso bei $(R-S)f$ — und bei RSf muß Sf und $R(Sf)$ Sinn haben. Bei R^* ist es sogar fraglich, ob ein solcher O. existiert. Für uns ist aber gerade die Operation $*$ von großer Wichtigkeit, wir werden darum gegebenenfalls gleich vom fertigen Paare R, R^* ausgehen, indem wir festsetzen: Ein konjugiertes O.-Paar (kurz: konj. O.-Paar) wird von zwei O. R, R^* mit demselben Definitionsbereich gebildet, für die (soweit beide Seiten sinnvoll sind) stets

$$(Rf, g) = (f, R^*g), \qquad (f, Rg) = (R^*f, g)$$

gilt (vgl. Anm. [12])). Außerdem verlangen wir, analog wie bei den H. O. und aus denselben Gründen (vgl. E. Einleitung V sowie Definition 6 und Anm. [30])), daß der Definitionsbereich die abg. lin. M. $\mathfrak{H}$ aufspanne. Diese Definition wird sich als die richtige Verallgemeinerung der Operation $*$ in $\mathfrak{B}$ erweisen, die H. O. sind in ihr als Spezialfall $R = R^*$ mit erfaßt.

2. Zunächst sei einiges über die Gesichtspunkte der Untersuchung von $\mathfrak{B}$ gesagt. Wir fassen diesen Ring als ein hyperkomplexes Zahlensystem auf: In der Tat ist es wohl das einfachste derartige System mit unendlich vielen Einheiten — d. h. das einfachste, welches über die im allgemeinen untersuchten und im wesentlichen durch die Gültigkeit der sogenannten „Teilerkettensätze" gekennzeichneten[13]) Systeme hinausgeht. Während aber in den hyperkomplexen Systemen mit endlich vielen Einheiten (bzw. den dies ersetzenden „Teilerkettensätzen") die Topologie so trivial ist, daß sie gar keine nennenswerte Rolle hat, spielt in unserem unendlichen System $\mathfrak{B}$ die Topologie eine entscheidende Rolle. So ist es z. B. bei der definitorischen Festlegung dessen, welche Teilmengen $\mathfrak{M}$ von $\mathfrak{B}$ auch Ringe (in der Terminologie der Anm. [13]) zitierten Arbeiten: „Ordnungen") sind — wenn wir nur verlangten, daß $\mathfrak{M}$ mit A, B auch aA, $A \pm B$, AB enthält, würden wir uns in ein unentwirrbares Dickicht pathologischer Bildungen verlieren —, es muß nämlich die Forderung gestellt werden, daß $\mathfrak{M}$ im Sinne einer (noch zu formulierenden) Topologie von $\mathfrak{B}$ abg. sei.

---

[13]) Vgl. E. Noether, Math. Annalen **96**, 1 (1926), S. 26–61, insbesondere § 2; ferner E. Artin, Abh. d. Math. Sem. d. Hamb. Univ. **5**, 3 (1927), S. 251–260.

Übrigens werden wir noch eine vereinfachende Forderung an die Ringe $\mathcal{M}$ stellen: mit A soll auch A^* dazugehören. Dies ist eine bedeutsame Einschränkung, die es ermöglicht, den Komplikationen der Elementarteilertheorie aus dem Wege zu gehen[14]). Es ist wohl motiviert, diese Schwierigkeiten zunächst auszuschalten; die Theorie von $\mathcal{B}$ liefert, gegenüber dem endlichvieldimensionalen Falle[15]), ohnehin genug neue Komplikationsmöglichkeiten.

Aus dem oben Gesagten geht hervor, daß wir uns zunächst um die Topologie von $\mathcal{B}$ kümmern müssen. Dabei wirkt es störend, daß es mehrere Möglichkeiten des Topologisierens (d. h. des Definierens des Umgebungsbegriffes) in $\mathcal{B}$ gibt, die für uns in Betracht kommen. Diese werden wir im § I genau diskutieren; es genüge hier, wenn wir sagen, daß bei Ringen $\mathcal{M}$ in $\mathcal{B}$ die Abg. der sogenannten schwachen Topologie verlangt werden soll.

Da der Durchschnitt beliebig vieler Ringe wieder ein Ring ist, ist insbesondere der Durchschnitt aller $\mathcal{M}$ umfassenden Ringe einer ($\mathcal{M}$ sei irgendeine Teilmenge von $\mathcal{B}$), er ist offenbar als kleinster $\mathcal{M}$ enthaltender Ring eindeutig gekennzeichnet — wir nennen ihn $\boldsymbol{R}(\mathcal{M})$. Eine weitere wichtige Begriffsbildung ist die folgende: Sei $\mathcal{M}$ eine Teilmenge von $\mathcal{B}$; die Menge aller A aus $\mathcal{B}$, für die A und A^* mit jedem B von $\mathcal{M}$ vert. sind[16]), heiße $\mathcal{M}'$. Diese Operation $'$ kann iteriert werden, dadurch entstehen aus $\mathcal{M}$ der Reihe nach die Mengen $\mathcal{M}', \mathcal{M}'', \mathcal{M}''', \ldots$. Sie hat eine Reihe einfacher Eigenschaften, so werden wir u. a. zeigen: Es ist stets $\mathcal{M} \subset \mathcal{M}''$,[17]) aus $\mathcal{M} \subset \mathcal{N}$ folgt $\mathcal{M}' \supset \mathcal{N}'$, es gilt allgemein $\mathcal{M}' = \mathcal{M}''' = \mathcal{M}^{\mathrm{V}} = \ldots$, $\mathcal{M}'' = \mathcal{M}^{\mathrm{IV}} = \mathcal{M}^{\mathrm{VI}} = \ldots$.

Die beiden Operationen $\boldsymbol{R}(\mathcal{M})$, $\mathcal{M}'$ ermöglichen zusammen eine einfache Kennzeichnung der algebraischen Verhältnisse in $\mathcal{B}$.

3. Ein O. A heißt normal, wenn A, A^* vert. sind (vgl. dazu E. Anhang I). Ein Ring $\mathcal{M}$ heißt Abelsch, wenn alle seine Elemente miteinander vert. sind. Man zeigt leicht: Ein Ring $\mathcal{M}$ ist dann und nur dann Abelsch, wenn alle seine Elemente normal sind, und der Ring $\boldsymbol{R}(A)$ ist es dann und nur dann, wenn A normal ist. Wir werden nun die Umkehrung beweisen: Jeder Abelsche Ring $\mathcal{M}$ kann durch einen normalen O. (sogar durch einen H. O.) erzeugt werden — d. h. es gibt ein solches A mit $\mathcal{M} = \boldsymbol{R}(A)$.

---

[14]) Für Matrizen in endlichvieldimensionalen Räumen hat E. Fischer diese Bedingung als erster mit Erfolg eingeführt.

[15]) Jedes hyperkomplexe System mit endlich vielen Einheiten kann ja durch Matrizen (also Operatoren) eines endlichvieldimensionalen Raumes dargestellt werden.

[16]) Zwei A, B aus $\mathcal{B}$ sind vert., wenn $AB = BA$ ist.

[17]) $\mathcal{M} \subset \mathcal{N}$ oder $\mathcal{N} \supset \mathcal{M}$ bedeutet: $\mathcal{M}$ ist Teilmenge von $\mathcal{N}$.

Eine weitere, bei Abelschen Ringen eintretende Vereinfachung ist die folgende: Sei $\mathfrak{M}$ eine Teilmenge von $\mathfrak{B}$, $\boldsymbol{R}(\mathfrak{M})$ der Ring von $\mathfrak{M}$ — man zeigt leicht, daß $\boldsymbol{R}(\mathfrak{M})$ entsteht, indem man zuerst aus den Elementen von $\mathfrak{M}$ alle durch (endlichvielmaliges) Anwenden der Operationen aA, $A+B$, AB und A^* entstehenden O. und deren Menge $\boldsymbol{r}(\mathfrak{M})$ bildet, und dann zu $\boldsymbol{r}(\mathfrak{M})$ alle seine Häufungspunkte hinzunimmt. Dabei ist Häufungspunkt, wie in 2. erwähnt, im Sinne der sogenannten schwachen Topologie zu verstehen. Wenn nun $\boldsymbol{R}(\mathfrak{M})$ Abelsch ist (dies ist dann und nur dann der Fall, wenn alle Elemente von $\mathfrak{M}$ miteinander und mit ihren * vert. sind), so werden wir zeigen, daß es genügt eine viel geringere Menge von Häufungspunkten — die sogenannten Limites stark doppelkonvergenter Folgen, vgl. die Diskussion dieser Begriffe im § I — zu $\boldsymbol{r}(\mathfrak{M})$ hinzuzufügen: schon dadurch entsteht ganz $\boldsymbol{R}(\mathfrak{M})$. Dieses Resultat ist insbesondere für die Theorie der Funktionen von normalen (also insbesondere H. und unitären) vert. O. wichtig, allerdings führt dieselbe aus dem Rahmen der vorliegenden Arbeit hinaus. —

Für beliebige (nicht notwendig Abelsche!) Ringe werden wir eine Beziehung zwischen $\boldsymbol{R}(\mathfrak{M})$ und $\mathfrak{M}''$ ($\mathfrak{M}$ eine beliebige Teilmenge von $\mathfrak{B}$) aufstellen. Wir werden insbesondere zeigen (unser weitestgehendes Resultat lautet noch etwas allgemeiner, vgl. § II, insbesondere Satz 5): Wenn 1 zu $\mathfrak{M}$ gehört, so ist $\boldsymbol{R}(\mathfrak{M})=\mathfrak{M}''$. Um die Bedeutung dieses Satzes für das als hyperkomplexes Zahlensystem aufzufassende $\mathfrak{B}$ einzusehen, vergegenwärtige man sich, was er besagt, wenn der Hilbertsche Raum $\mathfrak{H}$ durch einen endlichviel-, etwa k-dimensionalen Euklidischen Raum ersetzt wird. $\mathfrak{M}$ sei speziell eine (endliche oder auch kontinuierliche) Gruppe unitärer Matrizen, dann erkennt man sofort, daß $\boldsymbol{R}(\mathfrak{M})$ die Menge aller Linearaggregate von Matrizen aus $\mathfrak{M}$ ist (unter denen es ja höchstens k^2 lin. unabhängige gibt!).

Sei zunächst $\mathfrak{M}$ (als Matrizenmenge) irreduzibel, dies bedeutet bekanntlich, daß nur die Matrizen $a1$ mit allen Elementen von $\mathfrak{M}$ vert. sind[18]), d. h. daß $\mathfrak{M}'$ aus ihnen allein besteht. $\mathfrak{M}''$ umfaßt also alle Matrizen überhaupt, nach unserem Satze muß dies auch $\boldsymbol{R}(\mathfrak{M})$ tun — somit ist jede Matrix ein Linearaggregat der Matrizen aus $\mathfrak{M}$, oder was dasselbe ist (da es k^2 lin. unabhängige Matrizen gibt): es gibt in $\mathfrak{M}$ k^2 lin. unabhängige Matrizen. Oder auch: Keine feste lin. Relation zwischen den k^2 Matrizenelementen kann für alle Elemente von $\mathfrak{M}$ bestehen. Das ist aber der Burnsidesche Satz über irreduzible Matrizengruppen[19]). Wenn wir $\mathfrak{M}$ nicht

[18]) Vgl. I. Schur, Berl. Ber. 1905, S. 406. Die dort stets vorausgesetzte Endlichkeit von $\mathfrak{M}$ ist in diesem Falle bekanntlich unwesentlich.

[19]) Vgl. Burnside, Proc. London Math. Soc. (2) 3 (1905), S. 430. Allerdings nur für unitäre Matrizen!

als irreduzibel voraussetzen, erkennt man ebenso mühelos, daß unser Satz der Frobenius-I. Schurschen Verschärfung des Burnsideschen Satzes entspricht[20]).

Wir haben somit das Analogon desjenigen Satzes bewiesen, der in endlichvielen Dimensionen zur Begründung der unitären Darstellungstheorie von Gruppen dienen kann. Wieweit ein Äquivalent dazu im Hilbertschen Raume auf unseren Satz gegründet werden kann, soll hier nicht untersucht werden.

Mit diesen Ausführungen ist die erste Hälfte der vorliegenden Arbeit erschöpft.

4. Der Gegenstand der zweiten Hälfte (§§ IV—V) sind die am Ende von 1. erwähnten konj. O.-Paare — unabhängig von jeder Beschr.-Annahme. Zunächst gilt es, die normalen O. unter ihnen zu kennzeichnen. Hier entsteht nun eine Schwierigkeit: das Paar R, R^* sollte normal heißen, wenn R und R^* vert. sind — da aber R, R^* nicht überall sinnvoll sein müssen, ist die Definition der Produkte RR^*, R^*R bzw. die Zurückführung der Vert. auf dieselben zunächst unsicher[21]). Daß auch die naheliegende Matrizendefinition der Vert. versagt, werden wir im Anhang III sehen.

Wir helfen uns wie folgt. Wenn von zwei O. R, A wenigstens A überall sinnvoll ist, so definieren wir die Vert. so: AR sei Forts. von RA — d. h. wenn Rf Sinn hat, so ist auch $R(Af)$ sinnvoll, und zwar gleich $A(Rf)$ (Af, $A(Rf)$ sind ohnehin sinnvoll[22])). Daher können wir auch für eine Menge $\mathfrak{M}$ ganz beliebiger O. (die nicht einmal lin. und ebensowenig überall sinnvoll sein müssen) $\mathfrak{M}'$ bilden: Es ist die Menge aller A von $\mathfrak{B}$, für die A wie A^* mit jedem R von $\mathfrak{M}$ vert. ist — $\mathfrak{M}'$ ist also immer $\subset \mathfrak{B}$. Insbesondere haben wir für jeden O. R die Mengen $(R)'$, $(R)''$, …[23]) (alle $\subset \mathfrak{B}$!).

---

[20]) Vgl. Frobenius und I. Schur, Berl. Ber. 1906, S. 209.

[21]) Keinesfalls ist für die Vert. von R, S $RS = RS$ kennzeichnend. Sei nämlich Rf nicht überall sinnvoll, dann gilt von $0Rf$ dasselbe, während $R0f$ überall sinnvoll ist — also ist $0R \neq R0$, d. h. R mit 0 nicht vert., was bei einem vernünftigen Vert.-Begriff nicht eintreten darf! — Das Bilden von AB, $A \pm B$ bei nicht überall sinnvollen A, B ist überhaupt eine fragwürdige Angelegenheit. So gibt es z. B. hypermax. H. O. A, B, für die niemals (außer für $f = 0$) Af, Bf gleichzeitig sinnvoll sind, also $A + B$ gar nicht zu bilden ist. (Vgl. die demnächst im Journal f. Math. erscheinende Arbeit des Verf. „Zur Theorie der unbeschränkten Matrizen“, wo sehr allgemeine Klassen von Gegenbeispielen angegeben werden.)

[22]) Gerade das Beispiel $R, 0$ in Anm.[21]) legt diese Definition nahe. Wenn ferner E ein P. O. ist, so besagt die Vert. von R, E, daß R durch die zu E gehörige abg. lin. M. $\mathfrak{M}$ reduziert wird (vgl. E., Definition 13), wie es vernünftigerweise zu erwarten ist.

[23]) Es ist wohl nicht mißverständlich, daß wir, wie schon bei $R(A)$, die Menge mit dem einzigen Element R auch R nennen.

Nun zeigt man leicht (vgl. § II, 1.): Ein A von $\mathfrak{B}$ ist dann und nur dann normal, wenn $(A)''$ Abelsch ist. Diese Definition dehnen wir auf beliebige konj. O.-Paare aus: sie sollen normal heißen, wenn $(R)''$ Abelsch ist. Von der Anwendbarkeit dieser Definition auf weite Klassen von O. werden wir uns im Anhang II überzeugen: so ist die im folgenden zu entwickelnde Theorie ohne weiteres auf die (komplexen, unbeschränkten) Laurentschen Matrizen und ihre Verallgemeinerungen[24]) anwendbar.

Von diesen normalen O. werden wir nun zeigen, daß sie und nur sie eine (und zwar nur eine) Hilbertsche Spektralform mit komplexen Eigenwerten besitzen[25]). Diese Untersuchungen sind daher wesentlich anders gerichtet als die vorhergehenden Paragraphen und eher als eine Fortsetzung von E. anzusehen.

Wie man sieht, liegen bei normalen O. die Verhältnisse etwas günstiger wie bei H. O.: bei diesen war ja die Hilbertsche Spektralform nicht immer erreichbar[26]) — und dies ist um so sonderbarer, als im Endlichvieldimensionalen und auch noch im Beschränkten H. ein Spezialfall normal war ($AA^* = A^*A$ folgt aus $A = A^*$)! Dies klärt sich dadurch auf, daß wir zeigen werden: Ein H. O. ist im wesentlichen dann und nur dann normal, wenn er hypermax. ist, und gerade diese H. O. haben eine Hilbertsche Spektralform (vgl. E. Satz 36). Daß in den oben genannten Fällen H. als Spezialfall von normal erscheint, rührt also einfach daher, daß dort alles hypermax. ist.

5. Die Einteilung dieser Arbeit ist die folgende: Im § I werden die verschiedenen Topologisierungsmöglichkeiten von $\mathfrak{B}$ angegeben und durchdiskutiert. Im § II führen wir die Operationen $\mathfrak{M}'$ und $\boldsymbol{R}(\mathfrak{M})$ (Ringbildung) ein und beweisen verschiedene Eigenschaften derselben (u. a. das schon angekündigte Analogon des Burnside-Frobenius-Schurschen Satzes).

---

[24]) Vgl. Toeplitz, Math. Annalen **70** (1911), S. 351–376.

[25]) Zur Präzisierung dieses, die gewöhnliche (reelle) Hilbertsche Spektralform verallgemeinernden Begriffes vgl. E. Anhang II. — Der Begriff der Normalität bei endlichvieldimensionalen Matrizen stammt von Frobenius [Journ. f. Math. **84** (1877), S. 51–54]; er bewies, daß diese und nur diese Matrizen unitär auf die Diagonalform (mit komplexen Diagonalelementen!) transformiert werden können. Im Encyklopädie-Artikel von Hellinger und Toeplitz (II. C. 13, S. 1562–1563) wurde dasselbe für die vollstetigen O. des Hilbertschen Raumes gezeigt.

In E. Anhang I bewies der Verf., daß alle beschr. normalen O. auf die komplexe Hilbertsche Spektralform gebracht werden können (das ist eben das Äquivalent der Diagonalform). Hier soll die Theorie vervollständigt und auf unbeschr. O. ausgedehnt werden. — Unabhängig vom Verf., und unter Verwendung einer anderen Methode, hat seither A. Wintner die unitären O. (die ein Spezialfall der beschr. normalen sind) ebenfalls auf die Spektralform gebracht. Vgl. Math. Zeitschr. **30**, 1/2 (1929), S. 228–282.

[26]) Vgl. E., Einleitung VII.

Im § III diskutieren wir die Abelschen Ringe (vgl. das früher darüber Gesagte). Die §§ IV, V sind vom früheren (bis auf die Definitionen) fast unabhängig, ihr Inhalt ist die allgemeine Theorie der normalen Operatoren (ohne Voraussetzung der Realität oder der Beschränktheit) und deren Spektraldarstellung.

Anhang I behandelt die beschr. H. O., Anhang II gibt eine Anwendung der Theorie der normalen Operatoren auf Laurentsche und allgemeinere Matrizen. In Anhang III wird die Unbrauchbarkeit der üblichen Definition der Matrizenvertauschbarkeit im Unbeschränkten gezeigt.

I. Topologisierungen von $\mathfrak{H}$ und $\mathfrak{B}$.

1. Ehe wir die Topologie von $\mathfrak{B}$ betrachten, wollen wir die etwas einfacheren topologischen Verhältnisse in $\mathfrak{H}$ beschreiben — weniger der späteren Anwendungen halber, als um die Situation in $\mathfrak{B}$ exemplifizieren zu können.

Wir hatten bisher (wie in E.) nur die folgende Topologie in $\mathfrak{H}$ betrachtet, die als „starke" bezeichnet wird[27]), und die wir nach Hausdorff durch Angabe aller Umgebungen eines willkürlichen Punktes f_0 von $\mathfrak{H}$ kennzeichnen wollen:

Starke Topologie in $\mathfrak{H}$. Sei $U_1(f_0; \varepsilon)$ die Menge aller f mit $|f-f_0|<\varepsilon$; alle $U_1(f_0, \varepsilon)$ mit $\varepsilon>0$ sind die Umgebungen von f_0.

Diese Umgebungen sind somit die Kugelinneren um f_0. Daraus, daß $|f-f_0|$ eine Entfernung ist (vgl. E., Definition 4 und Anm. [27])) folgt sofort, daß die vier Hausdorffschen Umgebungsaxiome erfüllt sind (die man in jeder Topologie postuliert), und zwar[28]):

α) Ein Punkt ist in jeder seiner Umgebungen enthalten.

β) Der Durchschnitt von zwei Umgebungen eines Punktes umfaßt gewiß noch eine Umgebung desselben.

γ) Jeder Punkt in einer Umgebung besitzt selbst eine Umgebung, die Teilmenge der ersteren ist.

δ) Zwei verschiedene Punkte haben immer zwei elementfremde Umgebungen.

---

[27]) Vgl. Weyl, Dissertation Göttingen 1908, S. 8—9.

[28]) Vgl. Hausdorff, Grundzüge der Mengenlehre, I. Auflage, Leipzig und Berlin 1914; daselbst wird auf diesen Umgebungsbegriff die ganze Topologie aufgebaut. Wir erwähnen, wie Häufungspunkte und Limites zu definieren sind. Ein Punkt ist Häufungspunkt einer Menge, wenn jede seiner Umgebungen Punkte der Menge enthält; er ist Limes einer Folge, wenn jede seiner Umgebungen alle Punkte derselben, mit endlich vielen Ausnahmen, enthält. Man erkennt mühelos: Limes sein, heißt Häufungspunkt von jeder unendlichen Teilmenge der Folge sein.

Wir wissen ferner schon aus E., daß die Operationen af, $f \pm g$, (f, g), $|f|$ in diesem Sinne stetig sind (und zwar $f \pm g$, (f, g) in beiden Variablen gleichzeitig!).

In der starken Topologie gilt das sogenannte erste Hausdorffsche Abzählbarkeitsaxiom[29]): zu jedem f_0 gibt es eine absteigende Umgebungsfolge derart, daß jede Umgebung von f_0 eine solche aus dieser Folge umfaßt — man nehme die $U_1\left(f_0; \frac{1}{n}\right)$, $n = 1, 2, \ldots$. Hieraus folgert man mühelos: Zu jedem Häufungspunkte einer Menge (in $\mathfrak{H}$) gibt es eine Teilfolge dieser Menge, von der er der Limes ist.

Nun ist es auch üblich, eine andere Topologie in $\mathfrak{H}$ einzuführen, die sogenannte „schwache“[30]). D. h. gewöhnlich definiert man nur die schwachen Limites und nicht die Umgebungen, und zwar so: $f_1, f_2, \ldots$ konvergiert schwach gegen f, wenn für jedes feste φ $(f_n, \varphi) \rightarrow (f, \varphi)$. Es ist aber leicht, dies auf einen Umgebungsbegriff zurückzuführen:

Schwache Topologie in $\mathfrak{H}$. Sei $U_2(f_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ die Menge aller f mit $|(f - f_0, \varphi_1)| < \varepsilon, \ldots, |(f - f_0, \varphi_s)| < \varepsilon$; alle $U_2(f_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ mit beliebigen $\varphi_1, \ldots, \varphi_s$ (aus $\mathfrak{H}$, s beliebig) und $\varepsilon > 0$ sind die Umgebungen von f_0.

Wieder erkennt man leicht das Erfülltsein von Hausdorffs Axiomen α) bis δ), sowie die Tatsache, daß ein Limes im Sinne dieser Topologie mit dem oben angegebenen schwachen Limes identisch ist. Die Stetigkeit von af, $f \pm g$ (die letztere in beiden Variablen zugleich) ist evident, die von (f, g) in f — bei festem g — definitorisch, wegen $(f, g) = \overline{(g, f)}$ ist es also auch in g — bei festem f — so. Aber in beiden Variablen zugleich ist es nicht der Fall, denn dann wäre auch $|f| = \sqrt{(f, f)}$ stetig, was nicht zutrifft[31]).

Da $U_1(f_0; \varepsilon)$ Teilmenge von $U_2(f_0; \varphi_1, \ldots, \varphi_s, \eta)$ ist (z. B. für $\varepsilon = \eta : \mathrm{Max}(|\varphi_1|, \ldots, |\varphi_s|)$) ist jeder Häufungspunkt bzw. Limes im starken Sinne auch einer im schwachen — die Umkehrung kann schon wegen des verschiedenen Stetigkeitscharakters von $|f|$ nicht gelten.

Nun gilt in der schwachen Topologie von $\mathfrak{H}$ das erste Hausdorffsche Abzählbarkeitsaxiom nicht, denn seine wichtigste Folgerung ist nicht erfüllt: es gibt eine Menge und einen Häufungspunkt derselben, der für keine Teilfolge von ihr Limes ist — d. h. die schwache Abg. kann, im Gegensatze

---

[29]) Vgl. am in Anm. [28]) a. O., wegen der Separabilität von $\mathfrak{H}$ (E., § I, Axiom C) gilt auch das zweite, vgl. ebendort.

[30]) Weyl (vgl. Anm. [27])) definiert sie etwas anders als wir. Vgl. S. 9 a. a. O.

[31]) Denn in jedem $U_2(f_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ liegen noch f mit $|f| = 1$: es genügt f zu allen $\varphi_1, \ldots, \varphi_s$ orth. und mit dem Betrage 1 zu wählen!

zu dem, was man sonst gewohnt ist, nicht auf Konvergenz-Eigenschaften zurückgeführt werden.

Ehe wir das Gegenbeispiel angeben, erwähnen wir noch: wenn $f_1, f_2, \ldots$ schwach gegen f konvergiert, so sind die $|f_1|, |f_2|, \ldots$, nach einer Bemerkung von E. Schmidt, beschränkt[32]). Das Beispiel ist die Menge $\mathfrak{A}$ (wir denken uns $\mathfrak{H}$ als Folgenraum der $\{x_1, x_2, \ldots\}$ realisiert) aller $\{x_1, x_2, \ldots\}$ für die nur zwei $x_m, x_n \neq 0$ sind, und zwar $x_m = 1, x_n = m$ (sonst m, n beliebig). 0 ist schwacher Häufungspunkt von $\mathfrak{A}$, d. h. bei gegebenen $\{u_1^1, u_2^1, \ldots\}, \ldots, \{u_1^s, u_2^s, \ldots\}$, ε können wir $\left|\sum_{p=1}^{\infty} u_p^1 x_p\right| < \varepsilon, \ldots, \left|\sum_{p=1}^{\infty} u_p^s x_p\right| < \varepsilon$ ($\{x_1, x_2, \ldots\}$ aus $\mathfrak{A}$) erzwingen: wir wählen m so, daß $|u_m^1| < \frac{\varepsilon}{2}, \ldots, |u_m^s| < \frac{\varepsilon}{2}$ gilt (es ist ja $u_p^1 \to 0, \ldots, u_p^s \to 0$), und dann n so, daß $|u_n^1| < \frac{\varepsilon}{2m}, \ldots, |u_n^s| < \frac{\varepsilon}{2m}$ gilt. Dagegen konvergiert keine Teilfolge von $\mathfrak{A}$ schwach gegen 0: in einer solchen wären die Absolutwerte, d. i. $\sqrt{1 + m^2}$, beschränkt, d. h. es kämen nur endlich viele m darin vor — für ein m_0 müßte also

[32]) Dies folgt auch aus einem allgemeinen Satze von St. Banach, Fund. Math. 3 (1922), S. 157. Wir geben, der Vollständigkeit halber, seinen auf den vorliegenden Fall spezialisierten schönen Beweis wieder. Sei also $f_1, f_2, \ldots$ schwach gegen f konvergent; es gilt $|f_n| \leqq c$, mit festem c, zu beweisen. Es genügt $|(f_n, \varphi)| \leqq c\,|\varphi|$ (für alle φ!) zu zeigen: für $\varphi = f_n$ folgt daraus die Behauptung. Ferner gilt die obige Relation gewiß immer, wenn sie für die φ mit $|\varphi| = \varepsilon$ $(\varepsilon > 0)$ gilt, d. h. wenn $|(f_n, \varphi)|$ für alle n und alle $|\varphi| = \varepsilon$ beschränkt ist — also erst recht, wenn dies für alle $|\varphi| \leqq \varepsilon$ der Fall ist, d. h. in einer beliebigen Kugel um die 0. Wenn nun alle $|(f_n, \varphi)|$ beschränkt sind, falls φ in einer festen Kugel $\mathfrak{K}$ $(|\varphi - g_0| < \delta)$ variiert, sind sie auch in einer Kugel um die 0 beschränkt (nämlich $|\varphi| \leqq \delta$): denn ein solches φ ist gewiß Differenz von zwei Punkten von $\mathfrak{K}$ (z. B. $g_0 \pm \frac{1}{2}\varphi$). Wäre also der Satz falsch, so wären die $|(f_n, \varphi)|$ in jeder Kugel $\mathfrak{K}$ unbeschränkt: d. h. es gäbe zu jedem c ein φ_0 in $\mathfrak{K}$ und ein n_0, so daß $|(f_{n_0}, \varphi_0)| > c$ ist. Aus Stetigkeitsgründen muß dann $|(f_{n_0}, \varphi)| > c$ auch noch in einer geeigneten Kugel um φ_0 gelten, die noch als Teil von $\mathfrak{K}$ gewählt werden kann.

Und nun sei $\mathfrak{K}$ eine beliebige Kugel; $\mathfrak{K}'$ eine Kugel in $\mathfrak{K}$, in der stets $|(f_{n'}, \varphi)| > 1$ ist (n' fest), und zwar sei ihr Radius < 1; $\mathfrak{K}''$ eine Kugel in $\mathfrak{K}$, in der stets $|(f_{n''}, \varphi)| > 2$ ist (n'' fest), und zwar sei ihr Radius $< \frac{1}{2}, \ldots$. Da $\mathfrak{H}$ vollständig ist (vgl. E., § I, Axiom E) haben alle Kugeln $\mathfrak{K}, \mathfrak{K}', \mathfrak{K}'', \ldots$ einen gemeinsamen Punkt $\bar{\varphi}$, und für diesen ist $|(f_{n'}, \bar{\varphi})| > 1, |(f_{n''}, \bar{\varphi})| > 2, \ldots$, also $\limsup_{n \to \infty} |(f_n, \bar{\varphi})| = \infty$, entgegen der Annahme.

Wir bemerken noch: Sei $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System. Wenn $f_1, f_2, \ldots$ schwach gegen f konvergiert, so ist für jedes $\nu = 1, 2, \ldots$ $\lim_{n \to \infty} (f_n, \varphi_\nu) = (f, \varphi_\nu)$, ferner sind die $|f_n|$ beschränkt. Diese notwendige Bedingung ist auch hinreichend: die Menge der φ mit $\lim_{n \to \infty} (f_n, \varphi) = (f, \varphi)$ enthält die $\varphi_1, \varphi_2, \ldots$, sie ist offenbar eine lin. M., und weil die $|f_n|$ beschränkt (also die (f_n, φ) in φ gleichmäßig stetig) sind, ist sie abg. (im starken Sinne) — somit umfaßt sie ganz $\mathfrak{H}$. Damit ist alles bewiesen.

unendlich oft $m = m_0$ sein, d. h. $x_{m_0} = 1$, trotzdem wegen der schwachen Konvergenz nur endlich oft $x_{m_0} = 1$ sein dürfte. —

Zum Schluß sei noch auf den folgenden Umstand hingewiesen. Im Innern einer jeden festen Kugel $\mathfrak{K}$ sind alle f beschränkt, d. h. $|f| \leqq c$. Sei $\bar\varphi_1, \bar\varphi_2, \ldots$ ein vollst. norm. orth. System, dann enthält jedes $U_2(f_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ ein $U_2(f_0; \bar\varphi_1, \ldots, \bar\varphi_n, \delta)$ (soweit nämlich beide in $\mathfrak{K}$ liegen!): Sei nämlich $\varphi_1 = \sum_{\nu=1}^{\infty} a_\nu^1 \bar\varphi_\nu, \ldots, \varphi_s = \sum_{\nu=1}^{\infty} a_\nu^s \bar\varphi_\nu$, und n mit $\sum_{\nu=n+1}^{\infty} |a_\nu^1|^2, \ldots, \sum_{\nu=n+1}^{\infty} |a_\nu^s|^2 < \frac{\varepsilon^2}{16 c^2}$ gewählt, und weiter $\delta \leqq \varepsilon : 2 \operatorname{Max}\left(\sum_{\nu=1}^{n} |a_\nu^1|, \ldots, \sum_{\nu=1}^{n} |a_\nu^s|\right)$. Dann gilt für $r = 1, \ldots, s$ und alle f des letzteren U_2

$$|(f - f_0, \varphi_r)| \leqq \sum_{\nu=1}^{n} |a_\nu^r| \, |(f - f_0, \bar\varphi_\nu)| + \left(f - f_0, \sum_{\nu=n+1}^{\infty} a_\nu^r \bar\varphi_\nu\right)$$

$$\leqq \sum_{\nu=1}^{n} |a_\nu^r| \delta + |f - f_0| \sqrt{\sum_{\nu=n+1}^{\infty} |a_\nu^r|^2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d. h. sie gehören zum erstgenannten U_2. Da wir offenbar $\delta = \frac{1}{p}$ $(p = 1, 2, \ldots)$ wählen können, ist somit das erste Hausdorffsche Abzählbarkeitsaxiom erfüllt[33]).

Bekanntlich ist die schwache, nicht aber die starke, Topologie im Innern von $\mathfrak{K}$ sogar kompakt[34]). Also: die starke Topologie genügt stets dem Abzählbarkeitsaxiom, ist aber weder in ganz $\mathfrak{H}$ noch in $\mathfrak{K}$ kompakt — die schwache hat in ganz $\mathfrak{H}$ keine der beiden Eigenschaften, in $\mathfrak{K}$ dagegen beide.

2. Nach dieser vorhergehenden Orientierung wenden wir uns den (uns eigentlich interessierenden) Verhältnissen in $\mathfrak{B}$ zu.

Um die Konvergenz einer O.-Folge $A_1, A_2, \ldots$ gegen ein A zu definieren, verlangt man in der Regel entweder die starke Konvergenz aller $A_n f$ gegen Af, oder die schwache — das letztere bedeutet $(A_n f, g) \to (Af, g)$ für alle f, g. Wir haben also: entweder $|(A_n - A) f| \to 0$ für alle f, oder $|((A_n - A) f, g)| \to 0$ für alle f, g. Dies ist die starke bzw. schwache Konvergenz der O. in $\mathfrak{B}$. Wie in 1. führen wir diese Konvergenzbegriffe auf entsprechende Topologien in $\mathfrak{B}$ zurück. Wir haben dann:

Starke Topolopie in $\mathfrak{B}$. Sei $U_3(A_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ die Menge aller A von $\mathfrak{B}$ mit $|(A - A_0)\varphi_1| < \varepsilon, \ldots, |(A - A_0)\varphi_s| < \varepsilon$; alle $U_3(A_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$

---

[33]) Da es eine stark, also auch schwach, überall dichte Folge in $\mathfrak{H}$ gibt, zeigt man leicht die Gültigkeit auch des zweiten, vgl. Anm. [29]).

[34]) Alle Aussonderungsmethoden (bei Konvergenzbeweisen) im Hilbertschen Raume kommen darauf heraus.

mit beliebigen $\varphi_1, \ldots, \varphi_s$ (aus $\mathfrak{H}$, s beliebig) und $\varepsilon > 0$ sind die Umgebungen von A_0.

Schwache Topologie in $\mathfrak{B}$. Sei $U_4(A_0; \varphi_1, \psi_1, \ldots, \varphi_s, \psi_s, \varepsilon)$ die Menge aller A von $\mathfrak{B}$ mit $|((A - A_0)\varphi_1, \psi_1)| < \varepsilon, \ldots, |((A - A_0)\varphi_s, \psi_s)| < \varepsilon$; alle $U_4(A_0; \varphi_1, \psi_1, \ldots, \varphi_s, \psi_s, \varepsilon)$ mit beliebigen $\varphi_1, \psi_1, \ldots, \varphi_s, \psi_s$ (aus $\mathfrak{H}$, s beliebig) und $\varepsilon > 0$ sind die Umgebungen von A_0.

Vom Erfülltsein der Hausdorffschen Axiome α) bis δ) überzeugt man sich wiederum leicht, ebenso davon, daß diese Topologien auf die oben genannten Konvergenzbegriffe führen. Ferner ist es klar, daß die Operationen aA, $A + B$ in beiden Topologien stetig sind (die letztere in beiden Variablen zugleich). A^* ist in der schwachen stetig (wegen $((A - A_0)\varphi, \psi) = ((A^* - A_0)\psi, \varphi))$, in der starken nicht (wir übergehen die leicht angebbaren Gegenbeispiele). AB ist in beiden Topologien in jeder der beiden Variablen stetig, wenn die andere festgehalten wird: daß sie es stark in A ist, erkennt man, wenn man $\varphi_1, \ldots, \varphi_s$ durch $B\varphi_1, \ldots, B\varphi_s$ ersetzt; daß sie es stark in B ist, folgt aus der Stetigkeit des (festen) A; daß sie es schwach in A ist, folgt durch Ersetzen von $\varphi_1, \psi_1, \ldots, \varphi_s, \psi_s$ mit $B\varphi_1, \psi_1, \ldots, B\varphi_s, \psi_s$; daß sie es schwach in B ist, folgt aus dem vorigen sowie $AB = (B^* A^*)^*$. In beiden Variablen zugleich ist aber A, B in keiner von beiden stetig (auch hier sind Gegenbeispiele so einfach, daß wir sie übergehen).

Da $U_3(A_0; \varphi_1, \ldots, \psi_s, \varepsilon)$ Teilmenge von $U_4(A_0; \varphi_1, \psi_1, \ldots, \varphi_s, \psi_s, \eta)$ ist (z. B. für $\varepsilon = \eta : \mathrm{Max}(|\psi_1|, \ldots, |\psi_s|)$), ist jeder Häufungspunkt bzw. Limes im starken Sinne auch einer im schwachen — die Umkehrung kann schon wegen des verschiedenen Stetigkeitscharakters von A^* nicht gelten.

Das erste Abzählbarkeitsaxiom gilt in keiner dieser Topologien: denn für jede gibt es Mengen mit einem Häufungspunkte, der Limes keiner Teilfolge derselben ist — wir werden beides zeigen, indem wir eine Menge $\mathfrak{A}$ angeben, von der 0 starker Häufungspunkt ist, während keine Teilfolge es auch nur zum schwachen Limes hat.

Zunächst stellen wir aber wieder fest: wenn $A_1, A_2, \ldots$ schwach (also erst recht wenn es stark) gegen A konvergiert, so sind die $A_1, A_2, \ldots$ gleichmäßig beschränkt — d. h. es existiert ein festes c, so daß für alle n und f $|A_n f| \leqq c \cdot |f|$ gilt. Es genügt dies für schwache Konvergenz zu zeigen, wo man es analog zum entsprechenden Satze in $\mathfrak{H}$ beweist[35]). Hier-

[35]) Nach E., Satz 12β) genügt es $|(A_n f, g)| \leqq C \cdot |f|\,|g|$ zu zeigen. Dies beweisen wir mit der Methode von St. Banach genau so wie die analoge E. Schmidtsche Relation in Anm. [32]), nur daß die Funktion (f_n, φ) (von n und φ) durch die Funktion $(A_n f, g)$ (von n und f, g) zu ersetzen ist — und die Kugeln $\mathfrak{K}, \mathfrak{K}', \mathfrak{K}'', \ldots$ (für φ) durch Kugelpaare $\mathfrak{K}, \mathfrak{L}, \mathfrak{K}', \mathfrak{L}', \mathfrak{K}'', \mathfrak{L}'', \ldots$ (für f bzw. g).

aus folgt nebenbei, daß AB im Sinne der starken Konvergenz (nicht der Topologie!) in beiden Variablen zugleich stetig ist: aus $A_n \to A$, $B_n \to B$ folgt $A_n Bf \to ABf$ und wegen $B_n f - Bf \to 0$ sowie der gleichmäßigen Stetigkeit der A_n $A_n B_n f - A_n Bf \to 0$ — d. h. $A_n B_n f \to ABf$ also $A_n B_n \to AB$ (alle $\to$ im Sinne der starken Konvergenz von $\mathfrak{H}$ bzw. $\mathfrak{B}$). Nun geben wir das gewünschte Gegenbeispiel an.

$\bar{A}_{m,n}$ sei der folgende O. aus $\mathfrak{B}$ ($\mathfrak{H}$ sei wieder als Folgenraum der $\{x_1, x_2, \dots\}$ realisiert):

$$\bar{A}_{m,n}\{x_1, x_2, \dots\} = \{y_1, y_2, \dots\},$$

$$y_m = x_m, \quad y_n = m x_n, \quad y_p = 0 \quad \text{für} \quad p \neq m, n;$$

und $\mathfrak{A}$ die Menge aller $\bar{A}_{m,n}$. 0 ist starker Häufungspunkt von $\mathfrak{A}$, d. h. bei gegebenen $\{u_1^1, u_1^2, \dots\}, \dots, \{w_1^s, w_2^s, \dots\}$, ε können wir

$$|\bar{A}_{m,n}\{u_1^1, u_1^2, \dots\}| < \varepsilon, \dots, |A_{m,n}\{u_1^s, u_2^s, \dots\}| < \varepsilon,$$

also

$$\sqrt{(u_m^1)^2 + m^2(u_n^1)^2} < \varepsilon, \dots, \sqrt{(u_m^s)^2 + m^2(u_m^s)^2} < \varepsilon$$

erzwingen: wir wählen m so, daß $|u_m^1| < \frac{\varepsilon}{\sqrt{2}}, \dots, |u_m^s < \frac{\varepsilon}{\sqrt{2}}$ gilt (es ist ja $u_p^1 \to 0, \dots, u_p^s \to 0$), und dann n so, daß $|u_n^1| < \frac{\varepsilon}{\sqrt{2}\,m}, \dots, |u_n^s| < \frac{\varepsilon}{\sqrt{2}\,m}$ gilt. Dagegen konvergiert keine Teilfolge von $\mathfrak{A}$ schwach gegen 0: in einer solchen müßte gleichmäßig $|\bar{A}_{m,n} f| \leqq c \cdot |f|$ gelten, d. h. m beschränkt sein, es kämen also nur endlich viele m darin vor — für ein m_0 müßte also unendlich oft $m = m_0$ sein, d. h. $(\bar{A}_{m,n}\varphi, \psi) = 1$, wenn $\varphi = \psi = \{x_1, x_2, \dots\}$ mit $x_{m_0} = 1$, $x_p = 0$ für $p \neq m_0$ ist, trotzdem wegen der schwachen Konvergenz nur endlich oft $(\bar{A}_{m,n}\varphi, \psi) = 1$ sein dürfte.

Der Umstand, daß in der starken Topologie A^* unstetig ist, ist für unsere Zwecke manchmal störend. Wir führen daher die starke Doppelkonvergenz ein: sie besteht, wenn $A_1, A_2, \dots$ gegen A und $A_1^*, A_2^*, \dots$ gegen A^* konvergiert. Aus dem früher Gesagten folgt: der starken Doppelkonvergenz gegenüber sind die Operationen aA, A^*, $A + B$, AB stetig (die zwei letzten in beiden Variablen zugleich).

Zum Schluß zeigen wir: Wenn die $A_{m,n}$ stark gegen die A_m (m fest, $n \to \infty$) konvergieren, und die A_n gegen A ($n \to \infty$), und wenn alle $A_{m,n}$ gleichmäßig stetig sind (d. h. ein festes c mit $|A_{m,n} f| \leqq c \cdot |f|$ existiert), so konvergiert eine geeignete Teilfolge A_{M_p, N_p} stark gegen A ($p \to \infty$). Sei nämlich $f_1, f_2, \dots$ eine in $\mathfrak{H}$ (im starken Sinne, vgl. E., § 1, Axiom C) überall dichte Folge. Wir wählen N_p mit

$$|A_{N_p} f_1 - A f_1| < \frac{1}{p}, \ldots, |A_{N_p} f_p - A f_p| < \frac{1}{p}$$

(es ist ja stets $A_n f \rightarrow A f$), und M_p mit

$$|A_{M_p, N_p} f_1 - A_{N_p} f_1| < \frac{1}{p}, \ldots, |A_{M_p, N_p} f_p - A_{N_p} f_p| < \frac{1}{p}$$

(es ist ja stets $A_{m,n} f \rightarrow A_n f$) — somit gilt $A_{M_p, N_p} f \rightarrow A f$ für alle $f = f_1, f_2, \ldots$, d. h. in einer überall dichten Menge in $\mathfrak{H}$. Da aber alle A_{M_p, N_p} sowie A gleichmäßig stetig sind, gilt es dann für alle f, wie behauptet wurde. Für die starke Doppelkonvergenz gilt offenbar dasselbe.

3. Man kann $\mathfrak{B}$ noch auf eine dritte Art topologisieren. Man definiere nämlich die Konvergenz von $A_1, A_2, \ldots$ gegen A so: $|(A_n - A) f|$ soll für $n \rightarrow \infty$ und alle f gleichmäßig gegen 0 gehen — d. h. $\leqq c_n \cdot |f|$ sein, $c_n \rightarrow 0$ —, oder auch $|((A_n - A) f, g)|$ soll für $n \rightarrow \infty$ und alle f, g es tun — d. h. $\leqq c_n \cdot |f|\,|g|$ sein, $c_n \rightarrow 0$. Nach E. (Satz 12 β)) ist aber beides dasselbe, wir nennen es mit naheliegender Bezeichnung gleichmäßige Konvergenz: wie wir sehen ist es gleichgültig, ob wir die starke oder die schwache so verschärfen. Die zugehörige Topologie ist offenbar diese:

Gleichmäßige Topologie in $\mathfrak{B}$. Sei $U_5(A_0; \varepsilon)$ die Menge aller A von $\mathfrak{B}$ mit $|(A - A_0) f| < \varepsilon' < \varepsilon$ für alle f und irgendein festes ε'; alle $U_5(A_0; \varepsilon)$ mit $\varepsilon > 0$ sind die Umgebungen von A_0.

Daß die Hausdorffschen Axiome α) bis δ) erfüllt sind, ist klar, ebenso das Stimmen des hieraus folgenden Konvergenzbegriffes. Auch sieht man, daß die Operationen $aA, A^*, A + B, AB$ stetig sind (die zwei letzten in beiden Variablen zugleich)[36]).

Da $U_3(A_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ Teilmenge von $U_5(A_0; \varepsilon)$ ist, ist jeder gleichmäßige Häufungspunkt bzw. Limes auch ein solcher im starken (und somit erst recht im schwachen) Sinne[37]) — die Umkehrung gilt nicht, da das erste Abzählbarkeitsaxiom im Starken und im Schwachen nicht gilt, wohl aber, wie wir gleich sehen werden, im Gleichmäßigen.

Da wir uns unter den $U_5(A_0; \varepsilon)$ auf die $U_5\left(A_0; \frac{1}{n}\right)$, $n = 1, 2, \ldots$, beschränken können, gilt das erste Abzählbarkeitsaxiom. Übrigens entspringt diese Topologie einer Metrik des linearen Raumes $\mathfrak{B}$: wenn wir nämlich

---

[36]) A^* darum, weil aus $|Af| \leqq c \cdot |f|$ (für alle f) $|A^* f| \leqq c \cdot |f|$ folgt — da diese Aussagen mit $|(Af, g)| \leqq c \cdot |f|\,|g|$ bzw. $|(A^* f, g)| \leqq c \cdot |f|\,|g|$ (für alle f, g, vgl. E., Satz 12 β)) gleichwertig sind, und $(A^* f, g) = \overline{(Ag, f)}$ ist.

[37]) Aus der gleichmäßigen Konvergenz folgt, da * für sie stetig ist, auch die starke Doppelkonvergenz.

als Betrag von A die obere Schranke von $|Af|$ für $|f| = 1$ erklären, d. h. das kleinste c, für welches stets $|Af| \leqq c \cdot |f|$ ist — er heiße $|A|$ —, dann gelten erstens offenbar die folgenden Relationen:

$$|aA| = |a|\,|A|, \quad |A + B| \leqq |A| + |B|, \quad |AB| \leqq |A|\,|B|,$$

also ist $|A - B|$ eine Entfernung im linearen Raume $\mathfrak{B}$ (vgl. E., § I, Axiom B, ferner Definition 4 und Anm. [27]), und zweitens ist $\mathcal{U}_5(A_0, \varepsilon)$ die Menge aller A mit $|A - A_0| < \varepsilon$: d. h. die gleichmäßige Topologie entsteht aus dieser Metrik. —

Wie in $\mathfrak{H}$ so ist auch in $\mathfrak{B}$ innerhalb einer jeden festen Kugel $\mathcal{K}$ (wo alle A $|A| \leqq c$ genügen) der Charakter der starken und der schwachen Topologie wesentlich verändert: hier erfüllen beide das erste Abzählbarkeitsaxiom. Sei nämlich $\bar{\varphi}_1, \bar{\varphi}_2, \ldots$ ein vollst. norm. orth. System, dann enthält jedes $\mathcal{U}_4(A_0; \varphi_1, \ldots, \varphi_s, \varepsilon)$ bzw. $\mathcal{U}_5(A_0; \varphi_1, \psi_1, \ldots, \varphi_s, \psi_s, \varepsilon)$ ein $\mathcal{U}_4(A_0; \bar{\varphi}_1, \ldots, \bar{\varphi}_n, \delta)$ bzw. $\mathcal{U}_5(A_0; \bar{\varphi}_{k_1}, \bar{\varphi}_{l_1}, \ldots, \bar{\varphi}_{k_m}, \bar{\varphi}_{l_m}, \delta)$: Wir setzen im ersten Falle $\varphi_1 = \sum_{\nu=1}^{\infty} a_\nu^1 \bar{\varphi}_\nu, \ldots, \varphi_0 = \sum_{\nu=1}^{\infty} a_\nu^s \bar{\varphi}_\nu$, und wählen n mit

$$\sum_{\nu=n+1}^{\infty} |a_\nu^1|^2 < \frac{\varepsilon^2}{16\,c^2}, \ldots, \sum_{\nu=n+1}^{\infty} |a_\nu^s|^2 < \frac{\varepsilon^2}{16\,c^2}$$

sowie

$$\delta \leqq \varepsilon : 2\,\mathrm{Max}\Big(\sum_{\nu=1}^{\infty} |a_\nu^1|, \ldots, \sum_{\nu=1}^{\infty} |a_\nu^s|\Big).$$

Dann gilt für $r = 1, \ldots, s$ und alle A des letzteren $\mathcal{U}_4$

$$\begin{aligned} |(A - A_0)\varphi_r| &\leqq \sum_{\nu=1}^{n} |a_\nu^r|\,|(A - A_0)\bar{\varphi}_\nu| + \Big|(A - A_0) \sum_{\nu=n+1}^{\infty} a_\nu^r \bar{\varphi}_\nu\Big| \\ &\leqq \sum_{\nu=1}^{n} |a_\nu^r|\,\delta + 2\,c \sqrt{\sum_{\nu=n+1}^{\infty} |a_\nu^r|^2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

d. h. sie gehören zum erstgenannten $\mathcal{U}_4$. Im zweiten Falle setzen wir $\varphi_1 = \sum_{\nu=1}^{\infty} a_\nu^1 \bar{\varphi}_\nu$, $\psi_1 = \sum_{\nu=1}^{\infty} b_\nu^1 \bar{\varphi}_\nu, \ldots, \varphi_s = \sum_{\nu=1}^{\infty} a_\nu^s \bar{\varphi}_\nu$, $\psi_s = \sum_{\nu=1}^{\infty} b_\nu^s \bar{\varphi}_\nu$, und wählen n mit $\sum_{\nu=n+1}^{\infty} |a_\nu^1|^2 < \eta^2$, $\sum_{\nu=n+1}^{\infty} |b_\nu^1|^2 < \eta^2, \ldots, \sum_{\nu=n+1}^{\infty} |a_\nu^s|^2 < \eta^2$, $\sum_{\nu=n+1}^{\infty} |b_\nu^s|^2 < \eta^2$, wo $\eta > 0$ mit

$$4\,c\,\mathrm{Max}\,(|\varphi_1|, |\psi_1|, \ldots, |\varphi_s|, |\psi_s|)\,\eta + \eta^2 = \frac{\varepsilon}{2}$$

gewählt ist, und setzen ferner $\delta \leqq \varepsilon : 2\,\mathrm{Max}\Big(\Big(\sum_{\nu=1}^{n} |a_\nu^1|\Big)^2, \ldots, \Big(\sum_{\nu=1}^{n} |a_\nu^s|\Big)^2\Big)$. Schließlich sei $m = n^2$ und $k_1, l_1, \ldots, k_m, l_m$ alle Paare ϱ, σ $\varrho \leqq n$, $\sigma \leqq n$. Dann gilt für $r = 1, \ldots, s$ und alle A des letzteren $\mathcal{U}_5$

$$\Big|((A-A_0)\varphi_r,\psi_r)-\Big((A-A_0)\sum_{\nu=1}^{n}a_\nu^r\overline{\varphi}_\nu,\sum_{\nu=1}^{n}b_\nu^r\overline{\varphi}_\nu\Big)\Big|$$

$$=\Big|-\Big((A-A_0)\varphi_r,\sum_{\nu=n+1}^{\infty}b_\nu^r\overline{\varphi}_\nu\Big)-\Big((A-A_0)\sum_{\nu=n+1}^{\infty}a_\nu^r\overline{\varphi}_\nu,\psi_1\Big)$$

$$+\Big((A-A_0)\sum_{\nu=n+1}^{\infty}a_\nu^r\overline{\varphi}_\nu,\sum_{\nu=n+1}^{\infty}b_\nu^r\overline{\varphi}_\nu\Big)\Big|$$

$$\leqq 2c\cdot|\varphi_r|\cdot\sqrt{\sum_{\nu=n+1}^{\infty}|b_\nu^r|^2}+2c\cdot\sqrt{\sum_{\nu=n+1}^{\infty}|a_\nu^r|^2}\cdot|\psi_r|$$

$$+\sqrt{\sum_{\nu=n+1}^{\infty}|a_\nu^r|^2\sum_{\nu=n+1}^{\infty}|b_\nu^r|^2}$$

$$<4c\,\mathrm{Max}\,(|\varphi_r|,|\psi_r|)\,\eta+\eta^2\leqq\frac{\varepsilon}{2},$$

$$\Big|\Big((A-A_0)\sum_{\nu=1}^{n}a_\nu^r\overline{\varphi}_\nu,\sum_{\nu=1}^{n}b_\nu^r\overline{\varphi}_\nu\Big)\Big|\leqq\sum_{\mu,\nu=1}^{n}|a_\mu^r|\,|b_\nu^r|\cdot|((A-A_0)\overline{\varphi}_\mu,\overline{\varphi}_\nu)|$$

$$\leqq\sum_{\mu,\nu=1}^{n}|a_\mu^r|\,|b_\nu^r|\cdot\delta\leqq\frac{\varepsilon}{2},$$

also

$$|((A-A_0)\varphi_r,\psi_r)|\leqq\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

d. h. sie gehören zum erstgenannten $\mathfrak{U}_5$. Da wir offenbar $\delta=\frac{1}{p}$ $(p=1,2,\ldots)$ wählen können, ist damit die obige Behauptung erwiesen.

4. Zum Schluß bemerken wir: jede Teilmenge $\mathfrak{M}$ von $\mathfrak{B}$ hat eine abzählbare Teilmenge $\mathfrak{N}$, so daß jedes Element A von $\mathfrak{M}$ Limes einer stark doppelkonvergenten Teilfolge von $\mathfrak{N}$ ist[38]).

Sei $\mathfrak{M}_c$ die Menge aller A von $\mathfrak{M}$ mit $|A|<c$: es genügt die Behauptung für alle $\mathfrak{M}_1,\mathfrak{M}_2,\ldots$ zu beweisen (sei dann $\mathfrak{N}$ bzw. gleich $\mathfrak{N}_1,\mathfrak{N}_2,\ldots$), da wir $\mathfrak{N}=\mathfrak{N}_1+\mathfrak{N}_2+\ldots$ setzen können — somit dürfen wir von Anfang an voraussetzen, daß in $\mathfrak{M}$ durchweg $|A|<c$ ist.

Wir realisieren $\mathfrak{H}$ als Folgenraum (der $\{x_1,x_2,\ldots\}$ mit endlichem $\sum_{n=1}^{\infty}|x_n|^2$), dann hat A eine Matrix $\{a_{\mu,\nu}\}$, indem allgemein

---

[38]) Also ist $\mathfrak{N}$ in $\mathfrak{M}$ stark wie schwach überall dicht, wir können insbesondere $\mathfrak{M}=\mathfrak{B}$ setzen. Für die gleichmäßige Topologie ist dies gewiß falsch, da dort leicht eine Menge von Kontinuum vielen A angegeben werden kann, die paarweise die Entfernung 1 haben. — Wenn wir $\mathfrak{M}$ einer Kugel $\mathfrak{K}$ gleichsetzen, so können wir in $\mathfrak{K}$ aus der Gültigkeit des ersten Abzählbarkeitsaxioms (im starken wie im schwachen Sinne) leicht auch die des zweiten folgern (vgl. Anm. [30]), [33])). Wir erwähnen noch, daß die schwache Topologie in $\mathfrak{K}$ kompakt ist (vgl. Ende von 1. und Anm. [34])).

$$A\{x_1, x_2, \ldots\} = \{y_1, y_2, \ldots\},$$
$$y_\nu = \sum_{\mu=1}^{\infty} a_{\mu,\nu} x_\mu$$

(A ist ja lin. und stetig!) gilt. Sei nun $p = 1, 2, \ldots$ und $a_{\mu,\nu}^{(p)}$ ($\mu, \nu = 1, \ldots, p$) irgendwelche rationalen Zahlen, so daß für alle $\mu, \nu = 1, \ldots, p$ $|a_{\mu,\nu}^{(p)} - a_{\mu,\nu}| < \frac{1}{p}$ ist, wir definieren $A^{(p)}$ durch

$$A^{(p)}\{x_1, x_2, \ldots\} = \{y_1, y_2, \ldots\},$$
$$y_\nu = \begin{cases} \sum_{\mu=1}^{p} a_{\mu\nu}^{(p)} x_\mu, & \text{für } \nu = 1, \ldots, p, \\ 0 \quad, & \text{für die übrigen } \nu. \end{cases}$$

Für $p \geqq m$ haben wir also:

$$\begin{aligned} |A^{(p)}\varphi_m - A\varphi_m|^2 &= \left| \sum_{\mu=1}^{p} a_{m\mu}^{(p)} \varphi_\mu - \sum_{\mu=1}^{\infty} a_{m\mu} \varphi_\mu \right|^2 \\ &= \left| \sum_{\mu=1}^{p} (a_{m\mu}^{(p)} - a_{m\mu}) \varphi_\mu - \sum_{\mu=\nu+1}^{\infty} a_{m\mu} \varphi_\mu \right|^2 \\ &= \sum_{\mu=1}^{p} |a_{m\mu}^{(p)} - a_{m\mu}|^2 + \sum_{\mu=p+1}^{\infty} |a_{m\mu}|^2 < \frac{1}{p} + \sum_{\mu=p+1}^{\infty} |a_{m\mu}|^2. \end{aligned}$$

Für $p \to \infty$ strebt dies offenbar gegen 0,[39]) und dies gilt für alle $m = 1, 2, \ldots$. Wir haben also $A^{(p)}\varphi_m \to A\varphi_m$; ebenso zeigt man $A^{(p)*}\varphi_m \to A^*\varphi_m$.

Nun betrachten wir alle O. B aus $\mathfrak{B}$, deren Matrix $\{b_{\mu\nu}\}$ aus lauter rationalen Zahlen besteht, von denen überdies nur endlich viele $\neq 0$ sind — sie bilden offenbar eine abzählbare Menge $B_1, B_2, \ldots$. Die $A^{(p)}$ von vorhin sind solche B, wir haben also zu jedem A von $\mathfrak{B}$ eine Folge $B_{\nu_1}, B_{\nu_2}, \ldots$, so daß $B_{\nu_p} f \to Af$, $B_{\nu_p}^* f \to A^* f$ für alle $f = \varphi_1, \varphi_2, \ldots$.

Jetzt gehen wir zu $\mathfrak{M}$ über. Wir untersuchen für jedes Paar ν, p ($= 1, 2, \ldots$) ob es ein A von $\mathfrak{M}$ mit $|B_\nu f - Af| < \frac{1}{p}$, $|B_\nu^* f - A^* f| < \frac{1}{p}$, für $f = \varphi_1, \ldots, \varphi_p$, gibt — wenn es welche gibt, so wählen wir ein solches A aus und nennen es $A_{\nu,p}$. Die $A_{\nu,p}$ bilden eine abzählbare Teilmenge $\mathfrak{N}$ von $\mathfrak{M}$; es soll gezeigt werden, daß sie das Gewünschte leistet.

---

[39]) Weil $\sum_{\mu=1}^{\infty} |a_{m\mu}|^2$ konvergiert. Die Konvergenz aller Reihen $\sum_{m=1}^{\infty} |a_{mn}|^2$, $\sum_{n=1}^{\infty} |a_{mn}|^2$ sieht man so ein: Die zweite konvergiert, weil $A\{0, \ldots, 0, \underset{m}{1}, 0, \ldots\} = \{a_{m1}, a_{m2}, \ldots\}$ ist, die erste aber infolgedessen auch, wenn man statt A A^* betrachtet (wobei aus a_{nm} $\overline{a_{mn}}$ wird).

Sei A irgendein Element von $\mathfrak{B}$, $p = 1, 2, \ldots$. Wir bilden die vorhin genannte Folge $B_{\nu_1}, B_{\nu_2}, \ldots$; für diese streben alle $|B_{\nu_q} f - Af|$, $|B^*_{\nu_q} f - A^*_f|$ mit $q \to \infty$ gegen 0 ($f = \varphi_1, \varphi_2, \ldots$) — wenn wir also $n = \nu_q$ groß genug wählen, so ist $|B_n f - Af| < \frac{1}{p}$, $|B^*_n f - A^* f| < \frac{1}{p}$ für alle $f = \varphi_1, \ldots, \varphi_p$. Daher kann $A_{n,p}$ gebildet werden (weil z. B. A so ist), und für alle $f = \varphi_1, \ldots, \varphi_p$ gilt $|B_n f - A_{n,p} f| < \frac{1}{p}$, $|B^*_n f - A^*_{n,p} f| < \frac{1}{p}$, also $|A_{n,p} f - Af| < \frac{2}{p}$, $|A^*_{n,p} f - Af| < \frac{2}{p}$. Wir nennen $A_{n,p}$ (auch n hängt von p ab!) kurz $\bar{A}_p$, dann gilt offenbar $\bar{A}_p f \to Af$, $A^*_p f \to A^* f$ für alle $f = \varphi_1, \varphi_2, \ldots$. Also auch für alle Linearaggregate endlich vieler derselben, d. h. in einer in $\mathfrak{H}$ (stark) überall dichten Menge; und somit, da die $\bar{A}_p$ alle zu $\mathfrak{N}$, also zu $\mathfrak{M}$ gehören und darum (in p) gleichmäßig stetig sind, für alle f überhaupt. Die $\bar{A}_1, \bar{A}_2, \ldots$ sind somit stark doppelkonvergent gegen A — sie bilden aber eine Teilfolge von $\mathfrak{N}$.

II. Eigenschaften von $\boldsymbol{R}(\mathfrak{M})$ und $\mathfrak{M}'$.

1. Wir gehen nun an die Definition der schon in Einleitung 2. erwähnten Operationen $\boldsymbol{R}(\mathfrak{M})$ und $\mathfrak{M}'$ heran.

Definition 1. Eine Teilmenge $\mathfrak{M}$ von $\mathfrak{B}$ heißt ein Ring, wenn sie erstens mit A, B auch aA, A^*, $A + B$, AB enthält, und zweitens schwach abg. ist [40]).

Definition 2. Da der Durchschnitt beliebig vieler Ringe wieder ein solcher ist, ist insbesondere der Durchschnitt aller, eine gegebene Teilmenge $\mathfrak{M}$ von $\mathfrak{B}$ umfassender Ringe ein Ring — der kleinste, der $\mathfrak{M}$ umfaßt. Er heiße $\boldsymbol{R}(\mathfrak{M})$, der von $\mathfrak{M}$ erzeugte Ring.

Definition 3. Sei $\mathfrak{M}$ eine Teilmenge von $\mathfrak{B}$. Die Menge aller A von $\mathfrak{B}$, für die A und A^* mit jedem B von $\mathfrak{M}$ vert. ist (vgl. Anm. [16])), heiße $\mathfrak{M}'$. So können auch Iterierte $\mathfrak{M}''$, $\mathfrak{M}'''$, ... gebildet werden.

$\mathfrak{M}'$ ist stets ein Ring: daß es mit A, B auch aA, A^*, $A + B$, AB enthält, ist klar, aber es ist auch schwach abg. — denn wenn A ein schwacher Häufungspunkt von $\mathfrak{M}'$ ist und B zu $\mathfrak{M}$ gehört, so folgt daraus, daß für alle A' von $\mathfrak{M}'$ $A'B = BA'$, $A'^*B = BA'^*$ gilt und diese vier Ausdrücke alle in A (schwach, bei festem B) stetig sind, $AB = BA$, $A^*B = BA^*$.

Wenn A zu $\mathfrak{M}$ und B zu $\mathfrak{M}'$ gehört, so sind B, B^* mit A vert., also A, A^* mit B: daher gehört A zu $\mathfrak{M}''$. Also ist $\mathfrak{M} \subset \mathfrak{M}''$. Ferner ist es klar, daß aus $\mathfrak{M} \subset \mathfrak{N}$ $\mathfrak{M}' \supset \mathfrak{N}'$ folgt. Wenn wir dies auf $\mathfrak{M} \subset \mathfrak{M}''$

[40]) Man beachte: schwach abg. ist mehr wie stark abg.!

anwenden, wird $\mathfrak{M}' > \mathfrak{M}'''$; wenn wir aber darin bloß $\mathfrak{M}$ durch $\mathfrak{M}'$ ersetzen, so wird $\mathfrak{M}' < \mathfrak{M}'''$ — also ist $\mathfrak{M}' = \mathfrak{M}'''$. Ersetzen wir hierin $\mathfrak{M}$ durch $\mathfrak{M}', \mathfrak{M}'', \ldots$, so wird:

$$\mathfrak{M}' = \mathfrak{M}''' = \mathfrak{M}^{\mathrm{V}} = \ldots, \quad \mathfrak{M}'' = \mathfrak{M}^{\mathrm{IV}} = \mathfrak{M}^{\mathrm{VI}} = \ldots.$$

Da $\mathfrak{N}'$ stets ein Ring ist und offenbar die 1 enthält, gilt dasselbe für $\mathfrak{M}''$; wegen $\mathfrak{M} < \mathfrak{M}''$ ist also $(\mathfrak{M}, 1)$[41] $< \mathfrak{M}''$, d. h. $\boldsymbol{R}(\mathfrak{M}, 1) < \mathfrak{M}''$ (also auch $\boldsymbol{R}(\mathfrak{M}) < \mathfrak{M}''$). Wir werden übrigens bald zeigen (Satz 5), daß sogar $\boldsymbol{R}(\mathfrak{M}, 1) = \mathfrak{M}''$ gilt.

Wir nennen eine Menge $\mathfrak{M}$ Abelsch, wenn alle Elemente von ihr sowie deren * miteinander vert. sind — für Mengen, die mit A auch A^* enthalten, z. B. Ringe, genügt also die Vert. aller Elemente. Abelsch sein bedeutet offenbar $\mathfrak{M} < \mathfrak{M}'$; da hieraus $\mathfrak{M}'' < \mathfrak{M}'''$ folgt, ist mit $\mathfrak{M}$ auch $\mathfrak{M}''$ Abelsch; und wegen $\mathfrak{M} < \mathfrak{M}''$ gilt auch die Umkehrung. Aus $\mathfrak{M} < \boldsymbol{R}(\mathfrak{M}) < \mathfrak{M}''$ folgt, daß auch $\boldsymbol{R}(\mathfrak{M})$ mit diesen Mengen gleichzeitig Abelsch ist — also: der Abelsche Charakter ist für alle drei Mengen $\mathfrak{M}$, $\boldsymbol{R}(\mathfrak{M})$, $\mathfrak{M}''$ gleichwertig.

Sei $\mathfrak{M}$ ein Ring. Wenn er Abelsch ist, ist für jedes A von $\mathfrak{M}$ A mit A^* vert., d. h. A normal (vgl. Einleitung 2., bei Anm.[16]). Wenn er nicht Abelsch ist, so enthält er zwei unvert. A, B; wir setzen $A_1 = \frac{A+A^*}{2}$, $A_2 = \frac{A-A^*}{2i}$, $B_1 = \frac{B+B^*}{2}$, $B_2 = \frac{B-B^*}{2i}$ — die A_1, A_2, B_1, B_2 sind H. O. aus $\mathfrak{M}$ und gewiß nicht alle vert. (wegen $A = A_1 + i A_2$, $B = B_1 + i B_2$), also können wir A, B gleich als H. O. annehmen. Dann gehört $C = A + iB$ zu $\mathfrak{M}$ und ist mit $C^* = A - iB$ unvert., also nicht normal. Also: Ein Ring $\mathfrak{M}$ ist dann und nur dann Abelsch, wenn alle seine Elemente normal sind — ein beliebiges $\mathfrak{M}$ also, wenn dies für $\boldsymbol{R}(\mathfrak{M})$ gilt.

2. Jetzt sollen die Beziehungen der obigen Begriffsbildungen zu den P. O. und unitären O. erörtert werden.

Satz 1. $\mathfrak{M}$ sei ein Ring, A ein (beschr.) H. O., $E(\lambda)$ seine Z. d. E.[42]).

---

[41]) Die Vereinigungsmenge von $\mathfrak{M}$ mit der 1.

[42]) In E. (Definition 17) wurde eine Z. d. E. als eine Schar von P. O. $E(\lambda)$ $(-\infty < l < \infty)$ mit den folgenden Eigenschaften definiert:

α) Wenn $\lambda \leq \mu$ ist, so ist $E(\lambda)$ Teil von $E(\mu)$.

β) Wenn $\lambda \geq \lambda_0$, $\lambda \to \lambda_0$, so gilt für jedes f $E(\lambda) f \to E(\lambda_0) f$.

γ) Wenn $\lambda \to -\infty$ oder $\to +\infty$, so ist stets $E(\lambda) f \to 0$ bzw. $\to f$.

[Alle Konvergenzen, wie stets in E., stark. Wir würden jetzt für β) sagen: $E(\lambda)$ ist in der starken Topologie von $\mathfrak{B}$ eine nach rechts halbstetige Funktion von λ; und für γ): es kann für $\lambda = \pm\infty$, in der starken Topologie von $\mathfrak{B}$ stetig als 1 bzw. 0 definiert werden.] Die Z. d. E. $E(\lambda)$ gehört zum (nicht notwendig beschr.) H. O. A (vgl. E. Satz 36), wenn weiter gilt:

(Fortsetzung der Fußnote [42]) auf nächster Seite.)

A gehört dann und nur dann zu $\mathcal{M}$, wenn alle $E(\lambda)$, $\lambda < 0$, und $1 - E(\lambda)$, $\lambda \geqq 0$, zu $\mathcal{M}$ gehören. Wenn $\mathcal{M}$ die 1 enthält, können wir einfacher sagen: wenn alle $E(\lambda)$ zu $\mathcal{M}$ gehören.

Beweis. Die zweite Behauptung folgt aus der ersten; betrachten wir also diese.

Zunächst ist die Bedingung hinreichend. Es mögen nämlich alle $E(\lambda)$, $\lambda < 0$, und $1 - E(\lambda)$, $\lambda \geqq 0$, zu $\mathcal{M}$ gehören, und $-c \leqq \lambda \leqq c$ sei das in Anm. [42]) erwähnte Intervall, außerhalb dessen $E(\lambda)$ gleich 0 bzw. 1 ist. Sei K_ε eine Kette von Zahlen von $-c$ bis c: $-c = \lambda_0 < \lambda_1 < \dots < \lambda_{n-1} < \lambda_n = c$, mit einer Maschenweite $\leqq \varepsilon$, d. h. $\lambda_\nu - \lambda_{\nu-1} \leqq \varepsilon$ für $\nu = 1, 2, \dots, n$ $(\varepsilon > 0)$, und in der die 0 vorkommt — etwa $\lambda_m = 0$. Dann gehört

$$A^{K_\varepsilon} = \sum_{\nu=1}^{n} \lambda_\nu (E(\lambda_\nu) - E(\lambda_{\nu-1}))$$

$$= \sum_{\nu=1}^{m-1} \lambda_\nu (E(\lambda_\nu) - E(\lambda_{\nu-1})) + \sum_{\nu=m+1}^{n} \lambda_\nu ((1 - E(\lambda_{\nu-1})) - (1 - E(\lambda_\nu)))$$

auch zu $\mathcal{M}$, und es ist

$$(A^{K_\varepsilon} f, g) = \sum_{\nu=1}^{n} \lambda_\nu ((E(\lambda_\nu) f, g) - (E(\lambda_{\nu-1}) f, g)).$$

Spektralform. Af hat dann und nur dann Sinn, wenn

$$\int_{-\infty}^{\infty} \lambda^2 \, d\,|E(\lambda) f|^2$$

endlich ist (Stieltjessches Integral, da der Integrand stets $\geqq 0$ ist, und der Ausdruck hinter dem d, wie sich zeigen läßt, nie abnimmt, ist das Integral endlich [konvergent] oder $+\infty$ [eigentlich divergent] — nun soll das erstere zutreffen). Ist dies der Fall, so gilt für alle g

$$(Af, g) = \int_{-\infty}^{\infty} \lambda \, d(E(\lambda) f, g)$$

(das Integral konvergiert absolut).

In E. wurde festgestellt, daß die sogenannten hypermax. H. O. solche Z. d. E. haben, und nur diese — dabei entsprechen sich alle hypermax. H. O. und alle Z. d. E. eineindeutig. Dies war die Verallgemeinerung der Hilbertschen Spektralform ins Unbeschr. (vgl. E. Satz 36).

Für beschr. H. O. A ist die Z. d. E. stets vorhanden (dies ist das Resultat der Hilbertschen Theorie, vgl. auch E., wo sich zeigt, daß alle beschr. A hypermax. sind, Satz 48, α), β)). Für die Z. d. E. dieser H. O. ist es aber charakteristisch, daß sich die Eigenschaft γ) dahin verschärfen läßt, daß $E(\lambda)$ für genügend kleine bzw. große λ überhaupt konstant ist — also gleich 0 bzw. 1. Wenn dies etwa für $\lambda \leqq -c$ bzw. $\lambda \geqq c$ eintritt, so können wir also alle $\int_{-\infty}^{\infty}$ durch $\int_{-c}^{c}$ ersetzen (und γ) ist von selbst erfüllt).

Diese Kennzeichnung der Beschr. ist eines der Resultate der Hilbertschen Theorie, wir werden sie aber — lediglich der Vollständigkeit halber — im Anhang I kurz begründen.

Mit $\varepsilon \to 0$ strebt dies für alle f, g gegen $\int\limits_{-c}^{c} \lambda\, d(E(\lambda) f, g) = (Af, g)$, d. h. A ist schwacher Limes der A^{K_ε}.[43]) Also auch schwacher Häufungspunkt von $\mathfrak{M}$ — somit gehört es zu $\mathfrak{M}$.

Weiter ist die Bedingung notwendig. Es gehöre jetzt A zu $\mathfrak{M}$; dann gehören auch $A, A^2, A^3, \ldots$ dazu, und alle ihre Linearaggregate — d. h. alle $p(A)$, wo $p(x)$ jedes Polynom mit $p(0) = 0$ sein kann. Man beweist leicht:

$$(p(A) f, g) = \int\limits_{-c}^{c} p(\lambda)\, d(E(\lambda) f, g).\text{[44])}$$

Für $\lambda_0 < 0$ wählen wir ein $\varepsilon > 0$ und $p_\varepsilon(x)$ den Bedingungen

$$\left.\begin{array}{lll} 1-\varepsilon < p_\varepsilon(x) < 1+\varepsilon & \text{für} & -c \leqq x \leqq \lambda_0 \\ -\varepsilon < p_\varepsilon(x) < 1+\varepsilon & \text{für} & \lambda_0 \leqq x \leqq \lambda_0+\varepsilon, \\ -\varepsilon < p_\varepsilon(x) < \varepsilon & \text{für} & \lambda_0+\varepsilon \leqq x \leqq c, \end{array}\right\} \text{ sowie } p_\varepsilon(0) = 0$$

entsprechend (man erkennt mühelos, daß dies möglich ist). Dann ist für jedes f und g

$$(p_\varepsilon(A) f - E(\lambda_0) f, g) = \int\limits_{-c}^{c} p_\varepsilon(\lambda)\, d(E(\lambda) f, g) - \int\limits_{-c}^{\lambda_0} d(E(\lambda) f, g)$$

$$= \int\limits_{-c}^{\lambda_0} (p_\varepsilon(\lambda) - 1)\, d(E(\lambda) f, g) + \int\limits_{\lambda_0}^{\lambda_0+\varepsilon} p_\varepsilon(\lambda)\, d(E(\lambda) f, g) + \int\limits_{\lambda_0+\varepsilon}^{-c} p_\varepsilon(\lambda)\, d(E(\lambda) f, g).$$

In diesen drei Integralen sind die Integranden bzw. absolut $\leqq \varepsilon$, $1+\varepsilon$, ε, wir können daher mit Hilfe der Abschätzungsmethode aus Anm. [43]),

---

[43]) Es ist sogar gleichmäßiger Limes derselben. Denn wenn $f(\lambda)$ durch $f_\varepsilon(\lambda) = \lambda_\nu$ für $\lambda_{\nu-1} < \lambda \leqq \lambda_\nu$ definiert wird, so ist

$$(A^{K_\varepsilon} f, g) = \int\limits_{-c}^{c} f_\varepsilon(\lambda)\, d(E(\lambda) f, g), \qquad (Af, g) = \int\limits_{-c}^{c} \lambda\, d(E(\lambda) f, g),$$

$$((A^{K_\varepsilon} - A) f, g) = \int\limits_{-c}^{c} (f_\varepsilon(\lambda) - \lambda)\, d(E(\lambda) f, g).$$

Nun ist stets $|f_\varepsilon(\lambda) - \lambda| \leqq \varepsilon$, hieraus folgt wegen

$$|(E(\mu) f, g) - (E(\lambda) f, g)| \leqq \sqrt{|E(\mu) f|^2 - |E(\lambda) f|^2}\, \sqrt{|E(\mu) g|^2 - |E(\lambda) g|^2}$$

(vgl. die in E. beim Beweise von Satz 36) durchgeführte Abschätzung)

$$|((A^{K_\varepsilon} - A) f, g)| \leqq \varepsilon \cdot |f|\, |g|,$$

also (vgl. § I. 8.) $|A^{K_\varepsilon} - A| \leqq \varepsilon$. Damit ist alles bewiesen.

[44]) Die Rechnung entspricht in allen Einzelheiten der in E. Anhang I (beim Beweis von Satz 14*) für unitäre O. ausgeführten. Der Kunstgriff geht auf F. Riesz zurück.

$$|(p_\varepsilon(A)f - E(\lambda_0)f, g)|$$

$$\leqq \varepsilon\cdot|E(\lambda_0)f|\,|E(\lambda_0)g| + (1+\varepsilon)\cdot\sqrt{|E(\lambda_0+\varepsilon)f|^2 - |E(\lambda_0)f|^2}\sqrt{|E(\lambda_0+\varepsilon)g|^2 - |E(\lambda_0)g|^2}$$
$$+\varepsilon\cdot\sqrt{|f|^2 - |E(\lambda_0+\varepsilon)f|^2}\sqrt{|g|^2 - |E(\lambda_0+\varepsilon)g|^2} \;{}^{45})$$
$$\leqq \varepsilon\cdot|f|\,|g| + \sqrt{|E(\lambda_0+\varepsilon)f|^2 - |E(\lambda_0)f|^2}\sqrt{|E(\lambda_0+\varepsilon)g|^2 - |E(\lambda_0)g|^2}$$
$$\leqq \varepsilon\cdot|f|\,|g| + \sqrt{|E(\lambda_0+\varepsilon)f|^2 - |E(\lambda_0)f|^2}\,|g|$$

schließen. Da dies für alle g gilt, so folgt daraus[46])

$$|p_\varepsilon(A)f - E(\lambda_0)f| \leqq \varepsilon\cdot|f| + \sqrt{|E(\lambda_0+\varepsilon)f|^2 - |E(\lambda_0)f|^2},$$

was mit $\varepsilon \to 0$ gegen 0 strebt. Somit ist für $\varepsilon \to 0$ (etwa in der Folge $\varepsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots$) $p_\varepsilon(A)f \to E(\lambda_0)f$ (stark in $\mathfrak{H}$!) für jedes f — somit konvergieren die $p_\varepsilon(A)$ stark (in $\mathcal{B}$) gegen $E(\lambda_0)$. Also gehört $E(\lambda_0)$ zu $\mathcal{M}$.

Für $\lambda \geqq 0$ wählen wir Polynome $q_\varepsilon(x)$ mit

$$\left.\begin{array}{lll} -\varepsilon < q_\varepsilon(x) < \varepsilon & \text{für} & -c \leqq x \leqq \lambda_0, \\ -\varepsilon < q_\varepsilon(x) < 1+\varepsilon & \text{für} & \lambda_0 \leqq x \leqq \lambda_0+\varepsilon, \\ 1-\varepsilon < q_\varepsilon(x) < 1+\varepsilon & \text{für} & \lambda_0+\varepsilon \leqq x \leqq c, \end{array}\right\} \text{ sowie } q_\varepsilon(0) = 0$$

(auch dies ist ohne weiteres möglich), und erkennen dann durch ganz analoge Betrachtungen wie vorhin, daß die $q_\varepsilon(A)$ für $\varepsilon \to 0$ stark (in $\mathcal{B}$) gegen $1 - E(\lambda_0)$ konvergieren. In diesem Falle gehört also $1 - E(\lambda_0)$ zu $\mathcal{M}$.

Damit ist alles bewiesen.

Satz 2. Sei $\mathcal{M}$ ein Ring, $\mathcal{M}^P$, $\mathcal{M}^U$ die Menge aller P. O. bzw. aller unitären O. aus $\mathcal{M}$. Es ist stets $\boldsymbol{R}(\mathcal{M}^P) = \mathcal{M}$; $\mathcal{M}^U$ ist leer, wenn 1 nicht zu $\mathcal{M}$ gehört, im entgegengesetzten Falle aber ist $\boldsymbol{R}(\mathcal{M}^U) = \mathcal{M}$.

Beweis. Nach Satz 1 sind zwei Ringe $\mathcal{M}$, $\mathcal{N}$ identisch, wenn $\mathcal{M}^P = \mathcal{N}^P$ ist — denn sie enthalten dieselben H. O., also überhaupt dieselben O., da für die Zugehörigkeit eines beliebigen O. A die der zwei H. O. $\frac{A+A^*}{2}$, $\frac{A-A^*}{2i}$ charakteristisch ist. Nun ist $\mathcal{M}^P \subset \mathcal{M}$, $\boldsymbol{R}(\mathcal{M}^P) \subset \boldsymbol{R}(\mathcal{M}) = \mathcal{M}$, $(\boldsymbol{R}(\mathcal{M}^P))^P \subset \mathcal{M}^P$, und andererseits $\boldsymbol{R}(\mathcal{M}^P) \supset \mathcal{M}^P$, $(\boldsymbol{R}(\mathcal{M}^P))^P \supset \mathcal{M}^P$. Also gilt $(\boldsymbol{R}(\mathcal{M}^P))^P = \mathcal{M}^P$, d. h. $\boldsymbol{R}(\mathcal{M}^P) = \mathcal{M}$.

---

[45]) Wie in E., Beweis von Satz 36 (vgl. die vorhin zitierte Anm. [43])) wird hier die Ungleichheit

$$\sqrt{a_1}\sqrt{b_1} + \ldots + \sqrt{a_p}\sqrt{b_p} \leqq \sqrt{(a_1 + \ldots + a_p)(b_1 + \ldots + b_p)}$$

angewandt.

[46]) Man setze $g = p_\varepsilon(A)f - E(\lambda_0)f$.

Wenn U ein unitäres Element von $\mathcal{M}$ ist, so gehören auch U^* und $UU^* = 1$ zu $\mathcal{M}$; wenn also 1 nicht dazugehört, so ist $\mathcal{M}^U$ leer. Gehöre nun 1 zu $\mathcal{M}$. Erstens ist $\mathcal{M}^U \subset \mathcal{M}$, $\boldsymbol{R}(\mathcal{M}^U) \subset \boldsymbol{R}(\mathcal{M}) = \mathcal{M}$. Zweitens gehöre E zu $\mathcal{M}^P$, dann gehören $E, 1$ und somit $2E - 1$ zu $\mathcal{M}$; dieses ist aber unitär, weil E ein P. O. ist[47]), gehört also zu $\mathcal{M}^U$. $E = \frac{1}{2}((2E-1)+1)$ gehört daher zu $\boldsymbol{R}(\mathcal{M}^U)$, und hieraus folgt: $\mathcal{M}^P \subset \boldsymbol{R}(\mathcal{M}^U)$, $\boldsymbol{R}(\mathcal{M}^P) \subset \boldsymbol{R}(\mathcal{M}^U)$, also $\boldsymbol{R}(\mathcal{M}^U) \supset \mathcal{M}$. Damit ist $\boldsymbol{R}(\mathcal{M}^U) = \mathcal{M}$ bewiesen.

3. Sei $\mathcal{M}$ eine beliebige Teilmenge von $\mathcal{B}$. Wir betrachten alle f, für die bei jedem A von $\mathcal{M}$ $Af = 0$, $A^*f = 0$ gilt — diese bilden offenbar eine abg. lin. M. Der P. O. ihrer Komplementären (vgl. E. Definition 12) heiße E_0: Zugehörigkeit zu ihr wird also durch $E_0 f = 0$ gekennzeichnet. Wir definieren:

Definition 4. $\mathcal{M} \subset \mathcal{B}$ sei beliebig. Der soeben konstruierte P. O. E_0, für den $E_0 f = 0$ damit gleichwertig ist, daß für alle A von $\mathcal{M}$ $Af = 0$, $A^*f = 0$ gilt, heiße die Haupteinheit von $\mathcal{M}$. Wir zeigen zuerst:

Satz 3. Für alle A von $\mathcal{M}$, ja von $\boldsymbol{R}(\mathcal{M})$, gilt $E_0 A = AE_0 = A$.

Beweis. Gehöre A zu $\mathcal{M}$. Da identisch $E_0(1 - E_0)f = 0$ ist, ist $A(1 - E_0)f = 0$ — also $A(1 - E_0) = 0$, $A = AE_0$. Durch Anwenden auf A^* folgt ebenso $A^* = A^* E_0$, und wenn wir hierauf $*$ anwenden, $A = E_0 A$.

Betrachten wir nun alle A mit $E_0 A = AE_0 = A$. Sie bilden offenbar einen Ring, und dieser umfaßt, wie wir sahen, $\mathcal{M}$ — also ist er $\supset \boldsymbol{R}(\mathcal{M})$, womit alles bewiesen ist.

Satz 4. E_0 gehört zu $\mathcal{M}'$ und zu $\mathcal{M}''$.

Beweis. Daß alle A von $\mathcal{M}$ mit E_0 vert. sind, sahen wir schon, ferner ist $E_0^* = E_0$ — also gehört E_0 zu $\mathcal{M}'$. Gehöre nun B zu $\mathcal{M}'$. Wenn $E_0 f = 0$ ist, so ist für alle A von $\mathcal{M}$ $Af = 0$, also $BAf = ABf = 0$, und ebenso $A^*f = 0$, also $BA^*f = A^*Bf = 0$ — d. h. $E_0 Bf = 0$. Da identisch $E_0(1 - E_0)f = 0$ ist, haben wir also $E_0 B(1 - E_0)f = 0$, d. h. $E_0 B(1 - E_0) = 0$, $E_0 B = E_0 B E_0$. Wenden wir hierauf $*$ an und ersetzen wir B durch B^* (das ja auch zu $\mathcal{M}'$ gehört), so wird $BE_0 = E_0 B E_0$ — also $E_0 B = BE_0$. Dies gilt für alle B von $\mathcal{M}'$, also gehört E_0 auch zu $\mathcal{M}''$. —

Nun beweisen wir den schon in der Einleitung 3. angekündigten Satz.

Satz 5. $\boldsymbol{R}(\mathcal{M})$ ist die Menge aller A von $\mathcal{M}''$ mit $E_0 A = AE_0 = A$.

---

[47]) Es ist $(2E-1)^* = 2E-1$, $(2E-1)^2 = 4E^2 - 4E + 1 = 1$.

Beweis. Daß $\boldsymbol{R}(\mathcal{M})$ Teilmenge der genannten Menge ist, wissen wir schon; es ist nur zu zeigen, daß alle derartigen A wirklich zu $\boldsymbol{R}(\mathcal{M})$ gehören.

Seien $\varphi_1, \ldots, \varphi_k$ irgendwelche Elemente von $\mathfrak{H}$, die wir fürs Folgende zunächst festhalten, $k = 1, 2, \ldots$. Gleichzeitig denken wir uns noch einen Hilbertschen Raum $\bar{\mathfrak{H}}$ gegeben, und wir wählen in $\bar{\mathfrak{H}}$ k unendlichvieldimensionale abg. lin. M. $\bar{\mathfrak{H}}_1, \ldots, \bar{\mathfrak{H}}_k$ aus, die zueinander orth. sind und zusammen die abg. lin. M. $\bar{\mathfrak{H}}$ aufspannen[48]). Die $\bar{\mathfrak{H}}_1, \ldots, \bar{\mathfrak{H}}_k$ sind somit selbst Hilbertsche Räume, also mit $\mathfrak{H}$ isomorph[49]) — seien $\Gamma_1, \ldots, \Gamma_k$ isomorphe Abbildungen von $\mathfrak{H}$ auf bzw. $\bar{\mathfrak{H}}_1, \ldots, \bar{\mathfrak{H}}_k$.

Wir ordnen jetzt jedem A von $\mathcal{B}$ den Punkt $\Gamma_1(A\varphi_1) + \ldots + \Gamma_k(A\varphi_k)$ in $\bar{\mathfrak{H}}$ zu[50]), er heiße Bildpunkt von A. Sei $\bar{\mathfrak{L}}$ die Menge der Bildpunkte der A von $\boldsymbol{R}(\mathcal{M})$, offenbar eine lin. M., und $\bar{\bar{\mathfrak{M}}}$ ihre abg. Hülle (stark in $\bar{\mathfrak{H}}$!), also eine abg. lin. M. Der P. O. von $\bar{\bar{\mathfrak{M}}}$ heiße $\bar{E}$.

Wenn A ein beschr. O. in $\mathfrak{H}$ ist, so gibt es offenbar einen und nur einen beschr. O. $\bar{\bar{A}}$ in $\bar{\mathfrak{H}}$, für den

$$\bar{\bar{A}}(\Gamma_1 f_1 + \ldots + \Gamma_k f_k) = \Gamma_1(A f_1) + \ldots + \Gamma_k(A f_k)$$

gilt (vgl. Anm.[50])). Es ist offenbar $\bar{\bar{A}}^* = \overline{\overline{A^*}}$.

Nun gehöre B zu $\boldsymbol{R}(\mathcal{M})$. Wenn $\bar{f}$ zu $\bar{\mathfrak{L}}$ gehört, so ist $\bar{f} = \Gamma_1(A\varphi_1) + \ldots + \Gamma_k(A\varphi_k)$ für ein A von $\boldsymbol{R}(\mathcal{M})$. $\bar{\bar{B}} f = \Gamma_1(BA\varphi_1) + \ldots + \Gamma_k(BA\varphi_k)$, und da BA auch zu $\boldsymbol{R}(\mathcal{M})$ gehört, gehört $\bar{\bar{B}} f$ auch zu $\bar{\mathfrak{L}}$. $\bar{\bar{B}}$ bildet also $\bar{\mathfrak{L}}$ auf einen Teil von sich selbst ab — somit erst recht $\bar{\bar{\mathfrak{M}}}$. Jedes $\bar{E}\bar{f}$ liegt in $\bar{\bar{\mathfrak{M}}}$, also auch jedes $\bar{\bar{B}}\bar{E}\bar{f}$, somit gilt identisch $\bar{E}\bar{\bar{B}}\bar{E} f = \bar{\bar{B}}\bar{E} f$, d. h. $\bar{E}\bar{\bar{B}}\bar{E} = \bar{\bar{B}}\bar{E}$. Wir ersetzen B durch B^* (das ja auch zu $\boldsymbol{R}(\mathcal{M})$ gehört) und wenden * auf diese Gleichung an; wegen $\bar{\bar{B}}^* = \overline{\overline{B^*}}$ ist dann $\bar{E}\bar{\bar{B}}\bar{E} = \bar{E}\bar{\bar{B}}$; also $\bar{E}\bar{\bar{B}} = \bar{\bar{B}}\bar{E}$.

Fassen wir $\bar{E}$ näher ins Auge. Es ist offenbar

$$\bar{E}(\Gamma_1(f_1) + \ldots + \Gamma_k(f_k)) = \Gamma_1(g_1) + \ldots + \Gamma_k(g_k),$$

wobei die $g_1, \ldots, g_k$ lin. und stetig von den $f_1, \ldots, f_k$ abhängen. Also ist $g_l = E_{1l} f_1 + \ldots + E_{kl} f_k$ $(l = 1, \ldots, k)$, wobei die E_{jl} $(l, j = 1, \ldots, k)$

[48]) Zur Terminologie vgl. etwa E., § I. Die gewünschte Konstruktion erfolgt so: Sei $\bar{\varphi}_1, \bar{\varphi}_2, \ldots$ ein vollst. norm. orth. System, wir führen darin eine Doppelindizierung $\bar{\varphi}_{ln}$, $l = 1, \ldots, k$, $n = 1, 2, \ldots$, ein. Sei $\bar{\mathfrak{H}}_l$ die von $\bar{\varphi}_{l1}, \bar{\varphi}_{l2}, \ldots$ aufgespannte abg. lin. M. $(l = 1, \ldots, k)$ — diese $\bar{\mathfrak{H}}_1, \ldots, \bar{\mathfrak{H}}_k$ leisten alles, was wir verlangten.

[49]) Vgl. E., Einleitung III.

[50]) $\Gamma_1 f_1 + \ldots + \Gamma_k f_k$ durchläuft offenbar ganz $\bar{\mathfrak{H}}$, und zwar genau einmal, wenn $f_1, \ldots, f_k$ unabhängig voneinander $\mathfrak{H}$ durchlaufen.

lauter lin. stetige O. in $\mathfrak{H}$ sind[51]). Daher gilt

$$\bar E \bar{\bar B} f = \bar E \bar{\bar B}(\Gamma_1(f_1)+\dots+\Gamma_k(f_k)) = \bar E(\Gamma_1(Bf_1)+\dots+\Gamma_k(Bf_k))$$
$$= \Gamma_1(E_{11}Bf_1+\dots+E_{k1}Bf_k)+\dots+\Gamma_k(E_{1k}Bf_1+\dots+E_{kk}Bf_k),$$
$$\bar{\bar B}\bar E f = \bar{\bar B}\bar E(\Gamma_1(f_1)+\dots+\Gamma_k(f_k))$$
$$= \bar{\bar B}(\Gamma(E_{11}f_1+\dots+E_{k1}f_k)+\dots+\Gamma_k(E_{1k}f_1+\dots+E_{kk}f_k))$$
$$= \Gamma_1(BE_{11}f_1+\dots+BE_{k1}f_k)+\dots+\Gamma_k(BE_{1k}f_1+\dots+BE_{kk}f_k).$$

Also muß

$$E_{11}Bf_1+\dots+E_{k1}Bf_k = BE_{11}f_1+\dots+BE_{k1}f_k, \dots,$$
$$E_{1k}Bf_1+\dots+E_{kk}Bf_k = BE_{1k}f_1+\dots+BE_{kk}f_k$$

sein — d. h. allgemein $E_{jl}B = BE_{jl}$. Da B jedes Element von $\mathcal{M}$ sein kann, und dabei noch B, B^* zu $R(\mathcal{M})$ gehören, also mit E_{jl} vert. sind, gehören alle E_{jl} zu $\mathcal{M}'$.

Die Zugehörigkeit von $\bar f = \Gamma_1(f_1)+\dots+\Gamma_k(f_k)$ zu $\overline{\mathfrak{M}}$ ist durch $\bar E f = f$ charakterisiert, d. h.

$$\bar E(\Gamma_1(f_1)+\dots+\Gamma_k(f_k)) = \Gamma_1(f_1)+\dots+\Gamma_k(f_k),$$
$$\Gamma_1(E_{11}f_1+\dots+E_{k1}f_k)+\dots+\Gamma_k(E_{1k}f_1+\dots+E_{kk}f_k)$$
$$= \Gamma_1(f_1)+\dots+\Gamma_k(f_k),$$
$$(E_{11}-1)f_1+\dots+E_{k1}f_k = 0, \dots, E_{1k}f_1+\dots+(E_{kk}-1)f_k = 0.$$

Und wenn f der Bildpunkt von A ist, $f = \Gamma_1(A\varphi_1)+\dots+\Gamma_k(A\varphi_k)$, so haben wir die Bedingung:

$$E_{11}-1)A\varphi_1+\dots+E_{k1}A\varphi_k = 0, \dots, E_{1k}A\varphi_1+\dots+(E_{kk}-1)A\varphi_k = 0.$$

Gehört insbesondere A zu $\mathcal{M}''$, so daß es mit allen E_{jl} (da diese zu $\mathcal{M}'$ gehören) vert. ist, so wird hieraus

$$A[(E_{11}-1)\varphi_1+\dots+E_{k1}\varphi_k] = 0, \dots, A[E_{1k}\varphi_1+\dots+(E_{kk}-1)\varphi_k] = 0;$$

d. h. wir haben Bedingungen von der Form

$$A\omega_1 = 0, \dots, A\omega_k = 0,$$

mit festen $\omega_1, \dots, \omega_k$.

Für jedes A von $\mathcal{M}$ gehören A, A^* zu $R(\mathcal{M})$, also ihre Bildpunkte zu $\mathfrak{M}$, ferner gehören A, A^* auch zu $\mathcal{M}''$ — also ist $A\omega_1 = A^*\omega_1 = \dots = A\omega_k = A^*\omega_k = 0$. Daraus folgt (Definition 4) $E_0\omega_1 = \dots = E_0\omega_k = 0$. Wenn also A nur zu $\mathcal{M}''$ gehört, aber dabei $E_0A = AE_0 = A$ gilt, so haben wir

$$A\omega_1 = AE_0\omega_1 = 0, \dots, A\omega_k = AE_0\omega_k = 0,$$

---

[51]) Es ist, wie man leicht ausrechnet, $E_{jl} = \Gamma_j^{-1}P_{\bar{\mathfrak{H}}_j}\bar E\Gamma_l$, und umgekehrt $\bar E = \sum_{j,l=1}^{k}\Gamma_j E_{jl}\Gamma_l^{-1}P_{\bar{\mathfrak{H}}_l}$.

d. h. der Bildpunkt von A gehört zu $\overline{\mathfrak{M}}$. Somit gibt es für jedes $\varepsilon > 0$ ein A' von $\boldsymbol{R}(\mathcal{M})$, dessen Bildpunkt (das ist der allgemeine Punkt von $\overline{\mathfrak{L}}$) von demjenigen von A eine Entfernung $< \varepsilon$ hat, das hat aber wegen

$$\begin{aligned} &|\Gamma_1(A'\varphi_1) + \ldots + \Gamma_k(A'\varphi_k) - \Gamma_1(A\varphi_1) - \ldots - \Gamma_k(A\varphi_k)|^2 \\ &= |\Gamma_1((A'-A)\varphi_1) + \ldots + \Gamma_k((A'-A)\varphi_k)|^2 \\ &= |\Gamma_1((A'-A)\varphi_1)|^2 + \ldots + |\Gamma_k((A'-A)\varphi_k)|^2 \\ &= |(A'-A)\varphi_1|^2 + \ldots + |(A'-A)\varphi_k|^2 \end{aligned}$$

die Konsequenz $|(A'-A)\varphi_1| < \varepsilon, \ldots, |(A'-A)\varphi_k| < \varepsilon$. D. h.: A' gehört zur starken Umgebung $\mathcal{U}_4(A; \varphi_1, \ldots, \varphi_k, \varepsilon)$ von A.

Da nun $\varphi_1, \ldots, \varphi_k$ und $\varepsilon > 0$ ganz willkürlich waren, ist A starker Häufungspunkt von $\boldsymbol{R}(\mathcal{M})$, und da dieses als Ring sogar schwach abg. ist, gehört A dazu.

Damit ist alles bewiesen. —

Aus dem Beweise von Satz 5 läßt sich noch herauslesen:

Zusatz. Wenn eine Menge $\mathcal{M}$ von den Ringeigenschaften α), β) der Definition 1 wohl α) besitzt, an Stelle von β) aber nur die folgende, weniger weitgehende: wenn A, A^* starke Häufungspunkte von $\mathcal{M}$ sind, so gehört A zu $\mathcal{M}$ (dies ist weniger als die starke Abg., und erst recht als die schwache) — so ist sie denoch ein Ring.

Beweis. Beim Beweise von Satz 5 wurden von den Eigenschaften von $\boldsymbol{R}(\mathcal{M})$ nur die folgenden benützt: es ist $> \mathcal{M}$, $< \mathcal{M}''$, in ihm gilt stets $E_0 A = A E_0 = A$, mit A, B gehören auch $aA, A^*, A+B, AB$ dazu, und unter diesen Prämissen wurde gezeigt, daß jedes A von $\mathcal{M}''$ mit $E_0 A = A E_0 = A$ starker Häufungspunkt von $\boldsymbol{R}(\mathcal{M})$ ist (die starke Abg. von $\boldsymbol{R}(\mathcal{M})$, die wir nur brauchten, um hieraus zu folgern, daß A dazu gehört, soll jetzt nicht herangezogen werden). Bei unserem $\mathcal{M}$ können wir daher $\boldsymbol{R}(\mathcal{M})$ durch $\mathcal{M}$ selbst ersetzen, dann haben wir: jedes A von $\mathcal{M}''$ mit $AE_0 = E_0 A = A$ ist starker Häufungspunkt von $\mathcal{M}$. Da mit A auch A^* so ist, ist A^* auch starker Häufungspunkt von $\mathcal{M}$. Nach der (bisher unbenutzten) zweiten Hälfte der Annahme gehört also A zu $\mathcal{M}$.

Aus Satz 5 folgt daher $\boldsymbol{R}(\mathcal{M}) < \mathcal{M}$, d. h. $\boldsymbol{R}(\mathcal{M}) = \mathcal{M}$, also ist $\mathcal{M}$ ein Ring, wie behauptet wurde. —

Für Mengen mit der Ringeigenschaft α (Definition 1) ist also starke und schwache Abg. gleichwertig. Dasselbe zeigte in $\mathfrak{H}$ E. Schmidt für die lin. M.[59]), unser gegenwärtiges Resultat ist ein Analogon dazu in $\mathcal{B}$.

---

[59]) Vgl. Rend. d. Circ. Mat. d. Palermo 25 (1908), S. 57—73.

4. Wir haben soeben $\boldsymbol{R}(\mathcal{M})$ mit Hilfe von E_0 und $\mathcal{M}''$ gekennzeichnet ($\mathcal{M}$ sei wieder eine beliebige Teilmenge von $\mathcal{B}$), nun wollen wir $E_0, \mathcal{M}''$ auf $\boldsymbol{R}(\mathcal{M})$ zurückführen.

Satz 6. E_0 gehört zu $\boldsymbol{R}(\mathcal{M})$, und zwar ist es der größte P.O. in $\boldsymbol{R}(\mathcal{M})$, in dem Sinne, daß jeder P.O. aus $\boldsymbol{R}(\mathcal{M})$ Teil von E_0 ist.

Beweis. Da E_0 zu $\mathcal{M}''$ gehört und $E_0 E_0 = E_0$ ist, gehört es zu $\boldsymbol{R}(\mathcal{M})$ (Satz 5). Wenn E ein P.O. aus $\boldsymbol{R}(\mathcal{M})$ ist, so ist $EE_0 = E$, d. h. E Teil von E_0 (E., Satz 17).

Satz 7. $\mathcal{M}''$ ist die Menge aller $A + a \cdot 1$, oder auch die aller $A + a \cdot (1 - E_0)$ (A aus $\boldsymbol{R}(\mathcal{M})$, a komplex).

Beweis. Da E_0 zu $\mathcal{M}''$ gehört, bedeutet beides dasselbe; wegen $\boldsymbol{R}(\mathcal{M}) \subset \mathcal{M}''$, 1 aus $\mathcal{M}''$, ist es klar, daß die genannte Menge $\subset \mathcal{M}''$ ist — es bleibt zu zeigen, daß jedes B von $\mathcal{M}''$ die Form $A + a \cdot (1 - E_0)$, A aus $\boldsymbol{R}(\mathcal{M})$, hat.

Mit E_0, B gehört auch $E_0 B = B E_0 = A$ zu $\mathcal{M}''$ (da E_0 zu $\mathcal{M}'$, B zu $\mathcal{M}''$ gehört, sind sie vert.!), dabei ist $E_0 A = E_0^2 B = E_0 B = A$, $A E_0 = B E_0^2 = B E_0 = A$. Nach Satz 5 gehört also A zu $\boldsymbol{R}(\mathcal{M})$. $C = B - A$ gehört also auch zu $\mathcal{M}''$, und es ist $E_0 C = E_0 B - E_0 A = A - A = 0$, $C E_0 = B E_0 - A E_0 = A - A = 0$. Wenn wir hieraus $C = a \cdot (1 - E_0)$ schließen können, sind wir am Ziele.

Sei F ein P.O., der Teil von $1 - E_0$ ist. Dann ist F zu E_0 orth. (vgl. E., Satz 17), d. h. $E_0 F = F E_0 = 0$. Für jedes A von $\boldsymbol{R}(\mathcal{M})$ ist somit $FA = F \cdot E_0 A = 0$, $AF = AE_0 \cdot F = 0$, d. h. A, F vert. — also für jedes A von $\mathcal{M}$ A, A^* mit F vert., d. h. F gehört zu $\mathcal{M}'$. Somit sind F, C vert. Zusammen mit $E_0 C = C E_0 = 0$ folgt daraus das gewünschte $C = a(1 - E_0)$.[53]) —

Zum Schluß zeigen wir noch:

Satz 8. Die 1 enthaltenden Ringe, die $\mathcal{M}$ mit $\mathcal{M} = \mathcal{M}''$, und die Mengen $\mathcal{N}'$ sind untereinander identicch.

---

[53]) Dies zeigt man so: Sei $E_0 f = 0$, $f \neq 0$. Wie man sofort einsieht, gehört zu f nach $Fg = \frac{1}{|f|^2}(g, f) \cdot f$ ein P. O., und zwar folgt aus $E_0 f = 0$ $E_0 F = 0$, d. h. E_0, F sind orth. — also F, C vert. Somit ist $Cf = CFf = FCf = \frac{1}{|f|^2}(Cf, f) \cdot f = a_f \cdot f$. Nun wollen wir zeigen, daß die Zahl a_f für alle f dieselbe ist. $a_f = a_{cf}$ ($c \neq 0$) ist klar, d. h. $a_f = a_g$ für lin. abhängige f, g. Sind diese aber unabhängig, so haben wir $C(f+g) = a_{f+g} \cdot (f+g)$, $C(f+g) = Cf + Cg = a_f \cdot f + a_g \cdot g$, $(a_f - a_{f+g}) \cdot f + (a_g - a_{f+g}) \cdot g = 0$, also $a_f = a_g = a_{f+g}$. Also ist $a_f = a$ wirklich konstant, $Cf = af$ für $E_0 f = 0$ ($f \neq 0$ ist offenbar unwesentlich).

Da identisch $E_0(1 - E_0) f = 0$ ist, gilt für alle f $Cf = Cf - CE_0 f = C(1 - E_0) f$ $= a(1 - E_0) f$, also ist $C = a(1 - E_0)$.

Beweis. Wenn ein Ring $\mathfrak{M}$ die 1 enthält, so ist für seine Haupteinheit E_0 $E_0 = E_0 1 = 1$, und da für jedes A dann $E_0 A = 1A = A$, $AE_0 = A1 = A$ gilt, $\boldsymbol{R}(\mathfrak{M}) = \mathfrak{M}''$ (Satz 5), d. h. $\mathfrak{M} = \mathfrak{M}''$. Wenn ein $\mathfrak{M}$ der Bedingung $\mathfrak{M} = \mathfrak{M}''$ genügt, so ist $\mathfrak{M} = \mathfrak{N}'$ mit $\mathfrak{N} = \mathfrak{M}'$. Jede Menge $\mathfrak{N}'$ ist Ring und enthält die 1 (vgl. Anfang von 1.).

Damit ist die Gleichwertigkeit unserer drei Kriterien erwiesen.

III. Die Abelschen Ringe.

1. Daß unter Voraussetzung der Ringeigenschaft α) (Definition 4) schwache und starke Abg. dasselbe ist, sahen wir schon. Wir betrachten nun Abelsche Mengen, und wollen zeigen, daß bei diesen die Abg.-Eigenschaften noch weiter abgeschwächt werden dürfen.

Sei $\mathfrak{M}$ eine zunächst beliebige Menge $\subset \mathfrak{B}$, indem wir auf ihre Elemente die Operationen $aA, A^*, A+B, AB$ (beliebig, aber endlich, oft) anwenden, entsteht die Menge $\boldsymbol{r}(\mathfrak{M})$ — die offenbar $\subset \boldsymbol{R}(\mathfrak{M})$ und $\supset \mathfrak{M}$ ist. Wenn wir zu $\boldsymbol{r}(\mathfrak{M})$ alle A hinzufügen, für die A und A^* starke Häufungspunkte von $\boldsymbol{r}(\mathfrak{M})$ sind, entsteht so eine Menge $\boldsymbol{r}_1(\mathfrak{M})$, die $\subset \boldsymbol{R}(\mathfrak{M})$ und $\supset \mathfrak{M}$ ist, infolge der Symmetrie der Definition mit A auch A^* enthält, und mit A, B auch $aA, A+B, AB$ enthält (weil $\boldsymbol{r}(\mathfrak{M})$ es tut, und diese Operationen im Sinne der starken Topologie stetig sind — AB zwar nur in jeder Variablen für sich, fürs Vorliegende genügt aber dies, wie man leicht einsieht). Wenn B und B^* starke Häufungspunkte von $\boldsymbol{r}_1(\mathfrak{M})$ sind, so sind sie auch solche von $\boldsymbol{r}(\mathfrak{M})$ (denn $\boldsymbol{r}_1(\mathfrak{M})$ besteht aus Häufungspunkten von $\boldsymbol{r}(\mathfrak{M})$) — daher gehören sie zu $\boldsymbol{r}_1(\mathfrak{M})$.

Der Zusatz zu Satz 5 ist also anwendbar: $\boldsymbol{r}_1(\mathfrak{M})$ ist ein Ring, woraus $\boldsymbol{r}_1(\mathfrak{M}) = \boldsymbol{R}(\mathfrak{M})$ folgt. Durch Hinzufügen aller starken bzw. aller schwachen Häufungspunkte entstehe aus $\boldsymbol{r}(\mathfrak{M})$ bzw. $\boldsymbol{r}_2(\mathfrak{M})$, $\boldsymbol{r}_3(\mathfrak{M})$. Dann ist $\boldsymbol{r}_1(\mathfrak{M}) \subset \boldsymbol{r}_2(\mathfrak{M}) \subset \boldsymbol{r}_3(\mathfrak{M}) \subset \boldsymbol{R}(\mathfrak{M})$ klar, also $\boldsymbol{r}_1(\mathfrak{M}) = \boldsymbol{r}_2(\mathfrak{M}) = \boldsymbol{r}_3(\mathfrak{M}) = \boldsymbol{R}(\mathfrak{M})$.

Wenn nun $\mathfrak{M}$ Abelsch ist (d. h. alle Elemente von $\mathfrak{M}$ sowie deren * miteinander vert. sind), so genügt eine noch geringere Erweiterung von $\boldsymbol{r}(\mathfrak{M})$, um $\boldsymbol{R}(\mathfrak{M})$ zu erhalten, nämlich:

Satz 9. Wenn $\mathfrak{M}$ Abelsch ist, so entsteht $\boldsymbol{R}(\mathfrak{M})$ aus $\boldsymbol{r}(\mathfrak{M})$ bereits durch Hinzufügung aller Limites stark doppelkonvergenter Teilfolgen desselben[51]).

Beweis. Nennen wir diese Menge $\bar{\boldsymbol{r}}(\mathfrak{M})$, daß sie $\subset \boldsymbol{R}(\mathfrak{M})$ und $\supset \mathfrak{M}$ ist, ist klar, ebenso daß sie mit A, B auch $aA, A^*, A+B, AB$

[51]) Dies ist schon darum eine wesentliche Verschärfung, weil weder in der starken noch in der schwachen Topologie das erste Abzählbarkeitsaxiom gilt — d. h. jeder Häufungspunkt auch Limes ist (vgl. § I, 2.). — Ob dieser Satz auch unabhängig vom Abelschen Charakter von $\mathfrak{M}$ gilt, konnte nicht entschieden werden.

enthält (weil $\mathbf{r}(\mathcal{M})$ so ist, und diese Operationen im Sinne der starken Doppelkonvergenz stetig sind). Wenn also $\overline{\mathbf{r}}(\mathcal{M})$ stark abg. ist, so ist es ein Ring (Zusatz zu Satz 5), und daher $= \mathbf{R}(\mathcal{M})$. Nur die starke Abg. ist also zu beweisen[55]).

Wenn A starker Häufungspunkt von $\overline{\mathbf{r}}(\mathcal{M})$ ist, so ist es einer von $\mathbf{R}(\mathcal{M})$, gehört daher zu $\mathbf{R}(\mathcal{M})$ — und ist somit normal (§ I, 1.); bei beliebigem $\varphi_1, \ldots, \varphi_k, \varepsilon > 0$ gibt es ein A' aus $\overline{\mathbf{r}}(\mathcal{M})$ in $\mathcal{U}_s(A; \varphi_1, \ldots, \varphi_k, \varepsilon)$, d. h. mit

$$|(A' - A)\varphi_1| < \varepsilon, \ldots, |(A' - A)\varphi_k| < \varepsilon,$$

da A' zu $\overline{\mathbf{r}}(\mathcal{M})$ gehört, ist es auch normal. Obendrein sind A, A^*, A', A'^* als Elemente von $\mathbf{R}(\mathcal{M})$ alle vert. Wir setzen

$$\frac{1}{2}(A + A^*) = B, \quad \frac{1}{2i}(A - A^*) = C, \quad \frac{1}{2}(A' + A'^*) = B',$$

$$\frac{1}{2i}(A' - A'^*) = C',$$

dies sind dann lauter vert. (beschr.) H.O., und zwar B', C' aus $\mathbf{r}(\mathcal{M})$, ferner ist $A' - A = (B' - B) + i(C' - C)$.

Wenn nun B_1, C_1 zwei vert. (beschr.) H.O. sind, so ist

$$|(B_1 + iC_1)f|^2 = |B_1 f|^2 + (B_1 f, iC_1 f) + (iC_1 f, B_1 f) + |iC_1 f|^2$$
$$= |B_1 f|^2 + (-iC_1 B_1 f, f) + (iB_1 C_1 f, f) + |C_1 f|^2 = |B_1 f|^2 + |C_1 f|^2,$$

woraus auf Grund der früheren Ungleichheiten

$$|(B' - B)\varphi_1| < \varepsilon, \ldots, |(B' - B)\varphi_k| < \varepsilon$$

und

$$|(C' - C)\varphi_1| < \varepsilon, \ldots, |(C' - C)\varphi_k| < \varepsilon$$

folgt. Wir bezeichnen Max $(|B|, |C|)$ mit c und Max $(|B'|, |C'|)$ mit c', d. h. es ist stets $|Bf| \leqq c \cdot |f|$, $|Cf| \leqq c \cdot |f|$, $|B'f| \leqq c' \cdot |f|$, $|C'f| \leqq c' \cdot |f|$. Während c fest ist, hängt c' von $\varphi_1, \ldots, \varphi_k, \varepsilon$ ab — indem wir beide eventuell vergrößern, nehmen wir $c' \geqq c > 0$ an.

Jetzt soll das Verhältnis von B, B' näher untersucht werden. Sei $p(x)$ ein Polynom mit den folgenden Eigenschaften:

$$\left.\begin{array}{lll} -c < p(x) < \begin{cases} c \\ -c + \delta \end{cases} & \text{für} & -c' \leqq x \leqq -c, \\ \left.\begin{array}{r} x - \delta \\ -c \end{array}\right\} < p(x) < \begin{cases} x + \delta \\ c \end{cases} & \text{für} & -c \leqq x \leqq c, \\ \left.\begin{array}{r} -c \\ c - \delta \end{array}\right\} < p(x) < \quad c & \text{für} & c \leqq x \leqq c', \end{array}\right\}$$

[55]) Man beachte: es ist nicht einmal die Abg. von $\overline{\mathbf{r}}(\mathcal{M})$ im Sinne der starken Doppelkonvergenz selbstverständlich.

wobei wir über $\delta > 0$ noch verfügen werden. Man sieht: für $|x| \leqq c$ ist $|p(x) - x| < \delta$, für $|x| \leqq c'$ ist $|p(x)| < c$, für $|x|, |y| \leqq c'$ ist $|p(x) - p(y)| < |x - y| + 2\delta$ also $(p(x) - p(y))^2 - (x - y)^2 < 8c'\delta + \delta^2$ Wir wählen nun δ so, daß in der ersten Ungleichheit $|p(x) - x| < \varepsilon$ und in der letzten $(p(x) - p(y))^2 - (x - y)^2 < \varepsilon^2$ gilt ($p(x)$ hängt von c', δ, ε, also mittelbar und unmittelbar von $\varphi_1, \ldots, \varphi_k, \varepsilon$ ab).

Aus einem Satze von F. Riesz (vgl. E., Anhang II, Satz 3*, 4*) folgt für alle f $|p(B)f - Bf| \leqq \varepsilon \cdot |f|$, ebenso $|p(B')f| \leqq c \cdot |f|$. Ferner ist das Polynom $\varepsilon^2 + (x - y)^2 - (p(x) - p(y))^2 = q(x, y)$ im Quadrat $|x|, |y| \leqq c'$ stets $\geqq 0$, hieraus und aus der Vert. von B, B' sowie $|B|, |B'| \leqq c'$, kann man vermöge eines Analogons des soeben benutzten Satzes — auf welches wir gleich zurückkommen werden — folgern, daß der H.O. $q(B, B')$ definit ist. D. h.:

$$(q(B, B')f, f) \geqq 0, \quad ((p(B') - p(B))^2 f, f) \leqq ((B' - B)^2 f, f) + \varepsilon^2 (f, f),$$
$$|(p(B') - p(B))f|^2 \leqq |(B' - B)f|^2 + \varepsilon^2 |f|^2.$$

Für $f = \varphi_1, \ldots, \varphi_k$ haben wir also

$$|(p(B') - p(B))\varphi_l| \leqq \varepsilon \sqrt{1 + |\varphi_l|^2}.$$

Zusammen mit der früheren Ungleichheit ergibt das

$$|(p(B') - B)\varphi_l| \leqq \varepsilon\left(|\varphi_l| + \sqrt{1 + |\varphi_l|^2}\right).$$

Wenn wir also $\varepsilon = \eta : \underset{l=1,\ldots,k}{\operatorname{Max}}\left(|\varphi_l| + \sqrt{1 + |\varphi_l|^2}\right)$ setzen, $\eta > 0$ beliebig, so liegt $p(B')$ in $\mathcal{U}_3(B; \varphi_1, \ldots, \varphi_k, \eta)$. Dabei ist, wie wir sahen, $|p(B')| \leqq c$ — d. h. B ist starker Häufungspunkt des in der Kugel $|B''| \leqq c$ liegenden Teiles von $r(\mathcal{M})$, und zwar der H.O. darin. Dann ist es aber (§ I, 3.) auch Limes einer stark konvergenten Folge dortselbst, welche — da es sich ausschließlich um H.O. handelt — sogar doppelkonvergent ist. D. h.: B gehört zu $\overline{r}(\mathcal{M})$.

Genau so zeigt man, daß auch C dazu gehört, also auch $A = B + iC$, womit alles bewiesen ist.

Es bleibt noch übrig, den vorhin übergangenen Beweis der Verallgemeinerung des F. Rieszschen Satzes nachzuholen. Seien also B, B' zwei vert. H.O. mit $|B|, |B'| \leqq c'$, nach E., Anhang II, Satz 11* (man setze $A = B + iB'$) gibt es dann zwei Z. d. E. $E(\lambda)$, $F(\lambda)$, so daß stets

$$(Bf, g) = \int_{-c'}^{c'} \lambda\, d(E(\lambda)f, g), \qquad (B'f, g) = \int_{-c'}^{c'} \lambda\, d(F(\lambda)f, g)$$

ist, und alle $E(\lambda)$, $F(\mu)$ vert. sind. Wie dort setzen wir

$$\Delta(\lambda, \mu) = E(\lambda)\, F(\mu) = F(\mu)\, E(\lambda)$$

und merken uns, daß

$$\Delta(\lambda', \mu')\,\Delta(\lambda'', \mu'') = \Delta(\mathrm{Min}(\lambda', \lambda''), \mathrm{Min}(\mu', \mu''))$$

ist. Aus den obigen Formeln folgt (alle $\int\int$ sind über $\int\limits_{-c'}^{c'}\int\limits_{-c'}^{c'}$ zu erstrecken):

$$(Bf, g) = \int\int \lambda\, d(\Delta(\lambda, \mu) f, g), \qquad (B'f, g) = \int\int \mu\, d(\Delta(\lambda, \mu) f, g).$$

Genau wie beim Beweise von Satz 1 (vgl. dazu das Zitat von Anm. [41]) gilt dann

$$(q(B, B')f, g) = \int\int q(\lambda, \mu)\, d(\Delta(\lambda, \mu) f, g),$$

$$(q(B, B')f, f) = \int\int q(\lambda, \mu)\, d(\Delta(\lambda, \mu) f, f) = \int\int q(\lambda, \mu)\, d\,|\Delta(\lambda, \mu) f|^2.$$

Nun ist im Integrationsgebiet $q(\lambda, \mu) \geqq 0$, und es ist $\int\int\limits_{\mathfrak{R}} d\,|\Delta(\lambda, \mu) f|^2 \geqq 0$ für jedes Rechteck $\mathfrak{R}$ (vgl. E., Anhang II, § 5), also ist das letzte Integral rechts $\geqq 0$. Da demnach für alle f $(q(B, B')f, f) \geqq 0$ ist, ist $q(B, B')$ definit, wie behauptet wurde[56].

2. Ein Ring $\mathcal{M} = \boldsymbol{R}(A)$ mit einer Erzeugenden A ist dann und nur dann Abelsch, wenn es die Menge (A) ist, d. h. wenn A, A^* vert. sind: also für normales A. Wir wollen nun die schon in Einleitung 3 angekündigte Umkehrung beweisen: jeder Abelsche Ring (d. h. jeder, der nur normale Elemente hat) kann auf die Form $\boldsymbol{R}(A)$ gebracht werden, wo das normale A sogar H. gewählt werden kann.

Satz 10. Wenn A alle normalen O. durchläuft, so durchläuft $\boldsymbol{R}(A)$ alle Abelschen Ringe, und nur diese. Und zwar kann man bei gegebenem $\mathcal{M} = \boldsymbol{R}(A)$ A immer als H. O. wählen.

Beweis. Nach dem vorhin Gesagten genügt es, zu zeigen: zu jedem Abelschen Ring $\mathcal{M}$ existiert ein H. O. A mit $\mathcal{M} = \boldsymbol{R}(A)$.

Zunächst ist $\mathcal{M} = \boldsymbol{R}(\mathcal{M}^P)$ (Satz 2). Nach § I, 4. gibt es eine in $\mathcal{M}^P$ (stark) überall dichte abzählbare Menge $\mathcal{N} \subset \mathcal{M}^P$. Somit ist $\mathcal{M}^P \subset \boldsymbol{R}(\mathcal{N})$, $\mathcal{M} = \boldsymbol{R}(\mathcal{M}^P) \subset \boldsymbol{R}(\mathcal{N})$, und andererseits $\boldsymbol{R}(\mathcal{N}) \subset \boldsymbol{R}(\mathcal{M}^P) = \mathcal{M}$ — d. h. $\mathcal{M} = \boldsymbol{R}(\mathcal{N})$. Dabei ist $\mathcal{N}$ eine Folge von P. O. $E_1, E_2, \ldots$ aus $\mathcal{M}$ — also sind diese vert.

Im folgenden schreiben wir für E Teil von F kürzer $E \leqq F$, zwei P. O. E, F sind vergleichbar, wenn $E \leqq F$ oder $F \geqq E$ ist. Wir haben $\mathcal{M} = \boldsymbol{R}(E_1, E_2, \ldots)$ wo alle E_n vert. sind — sie sollen nun (unter Erhaltung der obigen Eigenschaft) so abgeändert werden, daß sie sogar ver-

---

[56]) Man beachte, daß zu diesem Beweise die Z. d. E. gegebener H. O. herangezogen werden mußten, während umgekehrt aus dem F. Rieszschen Satze deren Existenz bewiesen werden kann. — Über die hier in E. Anhang II verwendeten Stieltjes-Radonschen Doppelintegrale vgl. Wiener Akad. 122, 2 (1913), S. 1295–1438, insbesondere §§ I–II.

gleichbar sind. Genauer: wir werden eine Folge $F_1, F_2, \dots$ von P. O. mit den folgenden Eigenschaften angeben. $F_1 = E_1$; $F_2 \leqq F_1 \leqq F_3$ und F_1, F_2, F_3 entstehen durch die Operationen $+, -, \cdot$ aus E_1, E_2 und umgekehrt; $\dots$; $F_{2^{n-1}} \leqq F'_1 \leqq F_{2^{n-1}+1} \leqq F'_2 \leqq \dots \leqq F'_{2^{n-1}-2} \leqq F_{2^n-2} \leqq F'_{2^{n-1}-1} \leqq F_{2^n-1}$ (hier sind $F'_1, \dots, F'_{2^{n-1}-1}$ die $F_1, \dots, F_{2^{n-1}-1}$ der durch $\leqq$ festgelegten Größe nach geordnet) und $F_1, \dots, F_{2^n-1}$ entstehen durch die Operationen $+, -, \cdot$ aus $E_1, \dots, E_n$ und umgekehrt; .... Wenn wir solche F_p haben, so ist klar, daß jedes F_p durch $+, -, \cdot$ aus endlich vielen E_n entsteht, und umgekehrt, also ist $R(F_1, F_2, \dots) = R(E_1, E_2, \dots) = \mathfrak{M}$; und dabei sind offenbar alle F_p vergleichbar. Wir wollen sie nun herstellen.

Die Konstruktion verläuft induktiv: für $n = 1$ ist sie trivial, sie möge also für $n = m - 1$ schon geglückt sein, wir führen sie für $n = m$ aus. $F'_1, \dots, F'_{2^{m-1}-1}$ seien wieder die $F_1, \dots, F_{2^{m-1}-1}$ in der $\leqq$-Reihenfolge, wir setzen dann

$$F_{2^{m-1}} = E_m F'_1, \qquad F_{2^{m-1}+1} = F'_1 + E_m(F'_2 - F'_1),$$
$$F_{2^{m-1}+2} = F'_2 + E_m(F'_3 - F'_2), \dots,$$
$$F_{2^m-2} = F'_{2^{m-1}-2} + E_m(F'_{2^{m-1}-1} - F'_{2^{m-1}-2}),$$
$$F_{2^m-1} = F'_{2^{m-1}-1} + E_m(1 - F'_{2^{m-1}-1}).$$

Die $F_{2^{m-1}}, \dots, F_{2^m-1}$ sind hier mit Hilfe der $F'_1, \dots, F'_{2^{m-1}-1}, E_m$ ausgedrückt, d. h. mit Hilfe der $F_1, \dots, F_{2^{m-1}-1}, E_m$, also nach Induktionsannahme mit $E_1, \dots, E_{m-1}, E_m$; von den E_n ist nach derselben überhaupt nur E_m zu untersuchen, und da gilt

$$E_m = F_{2^{m-1}} + (F_{2^{m-1}+1} - F'_1) + \dots + (F_{2^m-1} - F'_{2^{m-1}-1}).$$

Schließlich ist auch $F_{2^{m-1}} \leqq F'_1 \leqq F_{2^{m-1}+1} \leqq F'_2 \leqq \dots \leqq F'_{2^{m-1}-2} \leqq F_{2^m-2} \leqq F'_{2^{m-1}-1} \leqq F_{2^m-1}$ leicht verifizierbar.

Betrachten wir nun die $F_1, F_2, \dots$, ihre Reihenfolge ist:

$$F_2 \leqq F_1 \leqq F_3, \qquad F_4 \leqq F_2 \leqq F_5 \leqq F_1 \leqq F_6 \leqq F_3 \leqq F_7,$$
$$F_8 \leqq F_4 \leqq F_9 \leqq F_2 \leqq F_{10} \leqq F_5 \leqq F_{11} \leqq F_1$$
$$\leqq F_{12} \leqq F_6 \leqq F_{13} \leqq F_3 \leqq F_{14} \leqq F_7 \leqq F_{15}, \dots .$$

Dieselbe kann wie folgt gekennzeichnet werden. Wir numerieren die F_p um, indem wir für F_p, $p = 2^{l_1} + 2^{l_2} + \dots + 2^{l_m}$ ($l_1 > l_2 > \dots > l_m \geqq 0$, dyadische Entwicklung!), F_ϱ schreiben, $\varrho = \frac{1}{2 \cdot 3^{l_1}} + \frac{2}{3^{l_1 - l_2}} + \dots + \frac{2}{3^{l_1 - l_m}}$, dann stehen sie in der Größen-Reihenfolge da: $\varrho < \sigma$ zieht $F_\varrho \leqq F_\sigma$ nach sich. Dabei durchläuft der Index ϱ, wie man sieht, die Menge aller Mittelpunkte der von der bekannten Cantorschen triadischen Menge ausgelassenen Inter-

valle[57]) — d. h. eine abzählbare Menge, die im Intervalle $0 < x < 1$ liegt, und keinen einzigen ihrer Häufungspunkte enthält.

Wir konstruieren nun eine Z. d. E. $G(\lambda)$ mit den folgenden Eigenschaften: für $\lambda < -1$ $G(\lambda) = 0$, für $\lambda \geqq 0$ $G(\lambda) = 1$, für $-1 \leqq \lambda < 0$ jedes $G(\lambda)$ (starker) Häufungspunkt der F_ϱ, insbesondere $G(\varrho - 1) = F_\varrho$. Dann ist die Menge $\mathcal{N}$ aller $G(\lambda)$, $\lambda < 0$, und $1 - G(\lambda)$, $\lambda \geqq 0$, $\subset \mathcal{M}$, $\boldsymbol{R}(\mathcal{N}) \subset \boldsymbol{R}(\mathcal{M}) = \mathcal{M}$, andererseits umfaßt sie alle F_ϱ, d. h. $F_1, F_2, \ldots$, so daß $\boldsymbol{R}(\mathcal{N}) \supset \boldsymbol{R}(F_1, F_2, \ldots) = \mathcal{M}$ ist — also $\boldsymbol{R}(\mathcal{N}) = \mathcal{M}$. Wenn also A der beschr. H. O. mit der Z. d. E. $G(\lambda)$ ist (vgl. Anhang I), so ist (Satz 1) $\boldsymbol{R}(A) = \boldsymbol{R}(\mathcal{N}) = \mathcal{M}$, und damit ist alles bewiesen. Daher gilt es, die Z. d. E. $G(\lambda)$ aufzustellen, und zwar genügt es, sie für $-1 \leqq \lambda < 0$ zu definieren.

Zunächst sei in jedem von der Cantorschen Menge ausgelassenen (offenen!) Intervalle (vgl. Anm. [57])), wenn dessen Mitte ϱ ist, $\bar{G}(\lambda - 1) = F_\varrho$. Damit ist $\bar{G}(\lambda)$ in einer in $-1 \leqq \lambda < 0$ überall dichten offenen Menge definiert — und zwar folgt aus $\lambda < \mu$ $\bar{G}(\lambda) \leqq \bar{G}(\mu)$, und aus $\lambda \to \lambda_0$, $\bar{G}(\lambda) f \to \bar{G}(\lambda_0) f$ ($\bar{G}(\lambda)$ ist in λ stark halbstetig!). Daher können wir seine Definition auf ganz $-1 \leqq \lambda < 0$ so ausdehnen: sei $\lambda_1 > \lambda_2 > \ldots$, $\lambda_n \to \lambda$ eine Folge, für die $\bar{G}(\lambda_1), \bar{G}(\lambda_2), \ldots$ schon Sinn haben, wegen $\bar{G}(\lambda_1) \geqq \bar{G}(\lambda_2) \geqq \ldots$ haben die $\bar{G}(\lambda_n)$ einen P. O. G zum starken Limes (vgl. E., Satz 19). G hängt nur von λ ab: denn für zwei solche Folgen $\lambda_1' > \lambda_2' > \ldots$, $\lambda_1'' > \lambda_2'' > \ldots$, können wir aus Teilfolgen eine neue $\lambda'_{m_1} > \lambda''_{n_1} > \lambda'_{m_2} > \lambda''_{n_2} > \ldots$ kombinieren, die auch konvergieren muß — also sind die Limites dieselben. Wenn $\bar{G}(\lambda)$ Sinn hat, so ist nach obigem $G = \bar{G}(\lambda)$ — wir bezeichnen allgemein G mit $G(\lambda)$.

Für $\mu < \lambda$ wählen wir $\lambda_1 > \lambda_2 > \ldots$, $\lambda_n \to \lambda$, $\mu_1 > \mu_2 > \ldots$, $\mu_n \to \mu$ mit $\mu_n < \lambda_n$, dann ist $\bar{G}(\mu_n) \leqq \bar{G}(\lambda_n)$, also aus Stetigkeitsgründen $G(\mu) \leqq G(\lambda)$. Ferner ist bei gegebenem f für geeignetes λ_n $|\bar{G}(\lambda_n) f|^2 \leqq |G(\lambda) f| + \varepsilon$, also für $\lambda \leqq \lambda' \leqq \lambda_n$ $(\lambda_n > \lambda)$ erst recht

$$|G(\lambda') f|^2 \leqq |G(\lambda_n) f|^2 = |\bar{G}(\lambda_n) f|^2 \leqq |G(\lambda) f|^2 + \varepsilon,$$
$$|G(\lambda') f - G(\lambda) f|^2 = |G(\lambda') f|^2 - |G(\lambda) f|^2 \leqq \varepsilon$$

(vgl. E., Satz 16) — d. h. für $\lambda' \geqq \lambda$, $\lambda' \to \lambda$ $G(\lambda') f \to G(\lambda) f$. Also ist $G(\lambda)$ die gewünschte Z. d. E., und wir sind am Ziele. —

---

[57]) Die genannte Menge umfaßt alle Zahlen $\sum_{s=1}^{\infty} \frac{\alpha_s}{3^s}$, wo alle $\alpha_s = 0, 2$ sind. Sie ist bekanntlich perfekt und nirgends dicht, und läßt die Intervalle

$$\sum_{s=1}^{k} \frac{\alpha_s}{3^s} + \frac{1}{3^{k+1}} < x < \sum_{s=1}^{k} \frac{\alpha_s}{3^s} + \frac{2}{3^{k+1}}$$

aus, die Mittelpunkte derselben sind die Punkte $\sum_{s=1}^{k} \frac{\alpha_s}{3^s} + \frac{1}{2 \cdot 3^k}$ — eine andere Schreibweise für unsere ϱ.

Wenn $\mathfrak{M}$ ein Abelscher Ring ist, so ist er nach dem soeben bewiesenen Satz 10 gleich $R(A)$, wobei A ein H. O, ist; und nach Satz 9 ist $R(A)$ die Menge der starken Doppellimites von $r(A)$. $r(A)$ aber ist (wegen $A = A^*$) im vorliegenden Falle die Menge aller $p(A)$, wo $p(x)$ alle Polynome mit $p(0) = 0$ durchläuft. Also: $\mathfrak{M}$ ist die Menge der Limites aller stark doppelkonvergenten Folgen $p_1(A), p_2(A), \ldots$ ($p_1(x)$, $p_2(x), \ldots$ für $x = 0$ verschwindende Polynome). Durch eine allgemeinere Fassung des Begriffes $f(A)$, als wir sie bisher benutzten (d. i. Ausdehnung auf unstetige $f(x)$), könnte man zeigen: $\mathfrak{M}$ ist die Menge aller $f(A)$ ($f(x)$ eine beliebige für $x = 0$ verschwinderde Funktion, für die die erweiterte Definition gilt). Wir wollen jedoch hier auf diese Dinge nicht eingehen.

IV. Allgemeine Eigenschaften der (unbeschränkten) normalen Operatoren.

1. Wir gehen zu unserem zweiten Gegenstande über, der Betrachtung unbeschränkter O., und der Definition der Normalität für solche. Daher müssen wir zunächst beliebige (nicht notwendig überall sinnvolle oder stetige — wir verlangen vorläufig nicht einmal Lin. und Abg.) O. betrachten, die wir $R, S, \ldots$ nennen werden (im Gegensatze zu den $A, B, \ldots$ aus $\mathfrak{B}$). Wir definieren:

Definition 5. aR ist derjenige O., der für die f Sinn hat, für die Rf sinnvoll ist, und zwar ist sein Wert dann $a \cdot Rf$; $R \pm S$ ist derjenige O., der für die f Sinn hat, für die Rf und Sf Sinn haben, und zwar ist sein Wert dann $Rf \pm Sf$; RS ist derjenige O., der für die f Sinn hat, für die Sf und $R(Sf)$ Sinn haben, und zwar ist sein Wert dann $R(Sf)$ [58]).

R, A sind vert. (es ist wesentlich, daß A zu $\mathfrak{B}$ gehört), wenn RA Fortsetzung von AR ist (d. h. bei sinnvollem Rf auch $R(Af)$ sinnvoll ist — $Af, A(Rf)$ sind es sowieso —, und zwar gleich $A(Rf)$) [59]).

---

[58]) Daß bei dieser Definition das kommutative und assoziative Gesetz der Addition sowie das assoziative der Multiplikation gelten, ist klar. Vom distributiven gilt immer $(R \pm S) \cdot T = R \cdot T \pm S \cdot T$, dagegen $R \cdot (S \pm T) = R \cdot S \pm R \cdot T$ nur für lin. und überall sinnvolles R (ist R nur lin., so ist die linke Seite Fortsetzung der rechten). Es ist $R + 0 = R$, $1 \cdot R = R \cdot 1 = R$, $R \cdot 0 = 0$ (wenn Rf für $f = 0$ verschwindet) dagegen $0 \cdot R$ nur dort definiert, wo R es ist. Analog gestaltet sich das Verhältnis von $+$ und $-$.

Man sieht: im Unbeschränkten sind die Operationen $+, -, \cdot$ nicht mehr so durchsichtig wie im Beschränkten.

[59]) Nach E., Einleitung V. ist R Fortsetzung von S, wenn Rf überall sinnvoll ist, wo Sf es ist, und dort beide gleich sind. — Wenn auch R überall sinnvoll ist (z. B. zu $\mathfrak{B}$ gehört), so bedeutet die jetzige Definition der Vert. $RA = AR$, d. h. die gewöhnliche Vert. Für beliebige R käme es kaum in Frage, $RA = AR$ zu verlangen, da dann R mit 0 nicht vert. wäre (vgl. Anm. [21]), [58])). Die Unsymmetrie unserer Definition (zwischen R, A) zeigt, daß eine direkte Definition der Vert., für beliebige R, S, auf Schwierigkeiten stoßen muß.

Nun können wir die Definition 3 auf Mengen $\mathfrak{M}$ beliebiger O. erweitern

Definition 3'. Wenn $\mathfrak{M}$ eine Menge beliebiger O. ist (nicht mehr notwendig $\subset \mathfrak{B}$!), so sei $\mathfrak{M}'$ die Menge aller A von $\mathfrak{B}$, für die A und A^* mit jedem R von $\mathfrak{M}$ vert. sind.

$\mathfrak{M}'$ ist also auf alle Fälle $\subset \mathfrak{B}$, $\mathfrak{M}''$, $\mathfrak{M}'''$, ... fallen somit stets unter die alte Definition. Was bleibt nun von den Eigenschaften von $\mathfrak{M}'$ noch erhalten?

Man sieht sofort: 1 gehört allenfalls zu $\mathfrak{M}'$, und mit A, B auch A^*, AB; wenn ferner alle Elemente von $\mathfrak{M}$ lin. sind, so gehören mit A, B auch aA, $A+B$ dazu. Seien nun alle Elemente von $\mathfrak{M}$ abg., A ein starker Häufungspunkt von $\mathfrak{M}'$ und R ein Element von $\mathfrak{M}$. Wenn Rf Sinn hat, so gibt es ein A'_n von $\mathfrak{M}'$ in $\mathfrak{U}_3\left(A;\, f,\, Rf,\, \frac{1}{n}\right)$, d. h. $A'_n f \to Af$, $RA'_n f = A'_n Rf \to ARf$, wegen der Abg. von R ist somit $R(Af)$ sinnvoll und gleich ARf. Somit sind A, R vert.

Wenn also A, A^* starke Häufungspunkte von $\mathfrak{M}'$ sind, so gehört A zu $\mathfrak{M}'$. Nunmehr nehmen wir zusammenfassend an, daß alle Elemente von $\mathfrak{M}$ abg. lin. sind. Dann sind nach dem vorher Gesagten alle Prämissen des Zusatzes zu Satz 5 erfüllt: also $\mathfrak{M}'$ ein Ring. Da 1 offenbar dazugehört, ist $\mathfrak{M}' = \mathfrak{M}'''$ (Satz 8).

Bei beliebigem $\mathfrak{M}$ können wir die Resultate vom Anfang von § II, 1 auf $\mathfrak{M}' \subset \mathfrak{B}$ anwenden: $\mathfrak{M}' \subset \mathfrak{M}''' = \mathfrak{M}^{\mathrm{V}} = \mathfrak{M}^{\mathrm{VII}} = \ldots$, $\mathfrak{M}'' = \mathfrak{M}^{\mathrm{IV}} = \mathfrak{M}^{\mathrm{VI}} = \ldots$, $\mathfrak{M}' = \mathfrak{M}'''$ dagegen können wir nicht schließen, da die Relation $\mathfrak{M} \subset \mathfrak{M}''$ fehlt (es ist ja $\mathfrak{M}'' \subset \mathfrak{B}$, aber nicht notwendig $\mathfrak{M} \subset \mathfrak{B}$!). Es gilt aber, wie wir sahen, wenn $\mathfrak{M}$ aus lauter abg. lin. O. besteht. —

Schließlich konstatieren wir, welche Elemente $\mathfrak{M}'^{\mathfrak{P}}$ und $\mathfrak{M}'^{\mathfrak{U}}$ haben. Zu $\mathfrak{M}'^{\mathfrak{P}}$ gehört jeder P. O. E, der mit allen R von $\mathfrak{M}$ vert. ist ($E = E^*$!) — d. h. jeder der alle R von $\mathfrak{M}$ reduziert (vgl. E., Definition 13). Zu $\mathfrak{M}'^{\mathfrak{U}}$ gehört jedes unitäre U, für welches U, $U^* = U^{-1}$ beide mit jedem R von $\mathfrak{M}$ vert. sind, d. h. wo RU Forts. von UR und RU^{-1} Forts. von $U^{-1}R$ ist. Man sieht sofort, daß dies bedeutet: $RU = UR$ (auch die Definitionsbereiche identisch!), oder auch: $U^{-1}RU = R$.

2. Wir führen nun den Begriff des konjugierten Paares (kurz: konj. Paar) von O. ein.

Definition 5. Zwei O. R, R^* bilden ein konj. Paar, wenn sie denselben Definitionsbereich haben, dieser die abg. lin. M. $\mathfrak{H}$ aufspannt, und in ihm durchweg

$$(Rf, g) = (f, R^* g), \qquad (f, Rg) = (R^* f, g)$$

(vgl. den Anfang der Anm. [19])) gilt.

Man sieht sofort: Mit R, R^* bilden auch R^*, R ein konj. Paar. Ein R kann nicht mit zwei verschiedenen R_1^*, R_2^* je ein konj. Paar bilden (für R_1^*, R_2^* sind die Definitionsbereiche gleichlautend vorgeschrieben, ebenso $(R_1^* f, g)$, $(R_2^* f, g)$ für eine die abg. lin. M. $\mathfrak{H}$ aufspannende Menge der g: also die $R_1^* f$, $R_2^* f$ selbst) — ebensowenig zwei verschiedene R_1, R_2 mit demselben R^* (vgl. die erste Bemerkung). Das Zeichen $*$ ist motiviert, wenn man beachtet, daß für ein A aus $\mathfrak{B}$ gerade A, A^* ein konj. Paar bilden. R ist dann und nur dann ein H. O., wenn R, R ein konj. Paar bilden.

Ein konj. Paar S, S^* ist Forts. eines anderen R, R^*, wenn der Definitionsbereich von S, S^* den von R, R^* umfaßt, und im letzteren stets $Rf = Sf$, $R^* f = S^* f$ gilt[60]). Wenn beide nicht identisch sind, sprechen wir (ganz wie bei H. O., vgl. E. § II) von eig. Forts.; ein konj. Paar ohne eig. Forts. ist max. (= maximal).

Wir bemerken noch: jede Forts. eines H. O. ist ein H. O., d. h. wenn R, R^* die Forts. S, S^* hat, so folgt aus $R = R^*$ $S = S^*$. In der Tat: seien Sf, Rg sinnvoll, dann ist

$$(Sf, g) = (f, S^* g) = (f, R^* g) = (f, Rg) = (f, Sg) = (S^* f, g),$$

und da diese g die abg. lin. M. $\mathfrak{H}$ aufspannen, $Sf = S^* f$ — also $S = S^*$.

Man erkennt ganz leicht (ganz wie bei H. O., vgl. E. § II, Satz 9), daß jedes konj. Paar zu einem solchen fortgesetzt werden kann, das aus zwei lin. O. besteht — dagegen ist es nicht ohne weiteres zu sehen, ob auch der abg. Charakter derselben so erreichbar ist[61]). Wir gehen hierauf nicht näher ein.

3. Ein A von $\mathfrak{B}$ ist normal, wenn A, A^* vert. sind, d. h. wenn (A) Abelsch ist, oder, was dasselbe bedeutet (vgl. § II, 1), wenn $(A)''$ Abelsch ist. Dies übertragen wir auf beliebige R (da stets $(R)'' \subset \mathfrak{B}$ ist!):

Definition 6. Sei R, R^* ein konj. Paar, es ist normal, wenn $(R)''$ (eine Menge $\subset \mathfrak{B}$!) Abelsch ist.

Auf Grund des vorhin Gesagten ist es klar, daß dies für O. aus $\mathfrak{B}$ die alte Definition ist.

Satz 11. R, R^* ist dann und nur dann normal, wenn R^*, R es ist.

Beweis: Wir zeigen $(R)' = (R^*)'$, hieraus folgt ja $(R)'' = (R^*)''$, und damit die Behauptung; ferner genügt es aus Symmetriegründen

---

[60]) Jede der letztgenannten Relationen folgt aus der anderen: denn das auf den Definitionsbereich der R, R^* eingeschränkte S, S^* ist noch immer ein konj. Paar.

[61]) Mit den genannten Methoden läßt sich bloß eine Art gemeinsame Abg. beider O. R, R^* des Paares erreichen: aus $f_n \to f$, $Rf_n \to f^*$, $R^* f_n \to f^{**}$ folgt die Sinnvollheit von Rf, $R^* f$, und daß diese bzw. gleich f^*, f^{**} sind.

$(R)' \subset (R^*)'$ zu beweisen, d. h. daß, wenn A, A^* mit R vert. sind, sie auch mit R^* vert. sein müssen.

Sei also R^*f sinnvoll, dann ist Rf, RAf, RA^*f es auch, also R^*Af, R^*A^*f gleichfalls. Wenn g ebenso ist, so haben wir:

$$(R^*Af, g) = (Af, Rg) = (f, A^*Rg) = (f, RA^*g) = (R^*f, A^*g) = (AR^*f, g),$$
$$(R^*A^*f, g) = (A^*f, Rg) = (f, ARg) = (f, RAg) = (R^*f, Ag) = (A^*R^*f, g),$$

und da diese g die abg. lin. M. $\mathfrak{H}$ aufspannen, $R^*Af = AR^*f$, $R^*A^*f = A^*R^*f$. Somit sind A, A^* mit R^* vert., wie behauptet wurde.

Satz 12. R, R (R ein H. O.!) ist, wenn R lin. abg. angenommen wird, dann und nur dann normal, wenn R hypermax. ist[62]).

Beweis: Zunächst stellen wir einige allgemeine Betrachtungen über abg. lin. H. O. an. Wir bilden die Cayleysche Transformierte U von R, $\mathfrak{E}$, $\mathfrak{F}$ seien ihr Definitionsbereich bzw. Wertvorrat (vgl. E. § V, Satz 2), beide abg. lin. M. Der P. O. von $\mathfrak{E}$ sei E (vgl. E § III, Satz 13).

Sei A mit R vert. Jedes φ mit sinnvollem $U\varphi$ hat die Form $Rf + if$, daher hat RAf Sinn, und es ist $A\varphi = ARf + iAf = RAf + iAf$. Also ist auch $U(A\varphi)$ sinnvoll, und zwar $= RAf - iAf = ARf - iAf = A(U\varphi)$, d. h. U, A sind vert. Sei nun umgekehrt A mit U vert. Jedes f mit sinnvollem Rf hat die Form $\varphi - U\varphi$, daher hat $UA\varphi$ Sinn, und es ist $Af = A\varphi - AU\varphi = A\varphi - UA\varphi$. Also ist auch $R(Af)$ sinnvoll, und zwar $= i(A\varphi + UA\varphi) = i(A\varphi + AU\varphi) = A(Rf)$, d. h. R, A sind vert. Aus alledem folgt aber $(R)' = (U)'$, und somit erst recht $(R)'' = (U)''$.

Jetzt zeigt sich die Hinreichendheit der Bedingung sofort: wenn R hypermax. ist, so ist U unitär (vgl. E., Satz 35), also aus $\mathfrak{B}$, ferner mit $U^* = U^{-1}$ vert. — also ist U, U^* normal, und mit ihm (wegen $(R)'' = (U)''$) auch R, R. Es bleibt übrig, die Notwendigkeit zu prüfen.

Sei also R, R normal. Gehöre A zu $(R)'$ — da es dann auch zu $(U)'$ gehört, ist mit Uf auch UAf sinnvoll, d. h. mit f gehört auch Af zu $\mathfrak{E}$. Ef gehört stets zu $\mathfrak{E}$, also auch AEf, also ist $EAEf = AEf$, — d. h. $EAE = AE$. Durch Anwenden von $*$ und Ersetzen von A durch A^* (das auch zu $(R)'$ gehört) folgt hieraus $EAE = EA$, also ist $AE = EA$, d. h. A, E vert. Somit ist A mit U, E vert., also auch mit UE — dies ist wichtig, weil $V = UE$ überall sinnvoll ist und so zu $\mathfrak{B}$ gehört. Da dies für jedes A von $(R)'$ (das mit A auch A^* enthält!) gilt, gehört V zu $(R)''$.

---

[62]) Abkürzung für hypermaximal, vgl. E., Definition 9. – Wenn R nicht lin. abg. ist, so betrachte man $\tilde{R}$ (E. § II, Satz 9). Da es H. ist, ist $\tilde{R}, \tilde{R}$ ein konj. Paar; da alles, was mit R vert. ist, es auch mit $\tilde{R}$ ist, ist $(R)' \subset (\tilde{R})'$, $(R)'' \supset (\tilde{R})''$, also auch $(\tilde{R})''$ Abelsch. Somit ist auch $\tilde{R}, \tilde{R}$ normal und daher unser Satz auf abg. lin. $\tilde{R}$ anwendbar.

Also auch V^*, und weil $(R)''$ Abelsch ist, sind V, V^* vert. Nun ist für alle f, g

$$(V^*Vf, g) = (Vf, Vg) = (U(Ef), U(Eg)) = (Ef, Eg) = (Ef, g),$$

also $V^*V = VV^* = E$, und auch dieses muß zu $(R)''$ gehören. Daher sind V, E vert., was die Konsequenz

$$EV = VE = UE \cdot E = UE = V$$

hat. Wenn φ zu $\mathfrak{E}$ gehört, so ist $\varphi = E\varphi$, $U\varphi = UE\varphi = V\varphi = EV\varphi$, d. h. auch $U\varphi$ gehört zu $\mathfrak{E}$, und daher auch $\varphi - U\varphi$. Aber die $\varphi - U\varphi$ müssen überall dicht liegen (E., Satz 24), also ist $\mathfrak{E}$ eine überall dichte abg. lin. M. — d. h. $\mathfrak{E} = \mathfrak{H}$, $E = 1$. Hieraus folgt:

$$V = U \cdot 1 = U, \quad V^*V = VV^* = 1, \quad \text{also} \quad U^*U = UU^* = 1,$$

d. h. U unitär, und daher R hypermax.

Wir weisen noch einmal auf die beim Beweise dieses Satzes nebenbei gefundene Tatsache hin, daß für einen hypermax. H. O. R und seine (unitäre) Cayleysche Transformierte $(R)' = (U)'$ gilt — und daher $(R)'' = (U)''$, $(R)''' = (U)'''$, ... $(U)''$ ist Abelsch, also $(U)'' \subset (U)''' = (U)'$, $(R)'' \subset (R)'$, und dabei $(R)' = (R)''' = \ldots$, $(R)'' = (R)^{\mathrm{IV}} = \ldots$ (weil es für U so ist, oder nach § IV, 1).

Satz 13. R sei (wie vorhin) ein hypermax. H. O., U seine Cayleysche Transformierte, $E(\lambda)$ seine Z. d. E., und $\mathcal{E}$ die Menge aller $E(\lambda)$. Dann ist $(R)' = (U)' = \mathcal{E}'$ — und daher

$$(R)'' = (U)'' = \mathcal{E}'', \quad (R)''' = (U)''' = \mathcal{E}''', \ldots .$$

Beweis: Aus der ersten Gleichung folgt alles, und da $(R)' = (U)'$ schon feststeht, ist nur $(U)' = \mathcal{E}'$ zu beweisen — was jedenfalls aus $\mathbf{R}(U) = \mathbf{R}(\mathcal{E})$ folgt[63]). Es genügt also zu zeigen: U gehört zu $\mathbf{R}(\mathcal{E})$, alle $E(x)$[64]) gehören zu $\mathbf{R}(U)$.

Daß U zu $\mathbf{R}(\mathcal{E})$ gehört, beweist man auf Grund der Relation

$$(Uf, g) = \int_0^1 e^{2\pi i x}\, d(E(x)f, g)$$

wörtlich, wie das Entsprechende in der ersten Hälfte des Beweises von Satz 1 für den dortigen beschr. H. O. A gezeigt wird. Um einzusehen, daß umgekehrt $E(x)$ zu $\mathbf{R}(U)$ gehört, vergegenwärtigen wir uns, wie diese $E(x)$ in E., Anhang II konstruiert wurden. Die $E(x)$ waren Differenzen

---

[63]) Allgemein gilt: $\mathcal{M} \subset \mathbf{R}(\mathcal{M}) \subset \mathcal{M}''$, $\mathcal{M}' \supset (\mathbf{R}(\mathcal{M}))' \supset \mathcal{M}''' = \mathcal{M}'$, $(\mathbf{R}(\mathcal{M}))' = \mathcal{M}'$, d. h. $\mathbf{R}(\mathcal{M})$ bestimmt $\mathcal{M}'$.

[64]) Wir schreiben $E(x)$, $0 < x < 1$, statt $E(\lambda)$, $-\infty < \lambda < \infty$, wobei der Zusammenhang $\lambda = -\operatorname{ctg} \pi x$, $x = -\frac{1}{\pi} \operatorname{arc\,ctg} \lambda$ (vgl. E., Beweis von Satz 36) besteht.

starker Limites gewisser $F(\varrho)$ $(0<\varrho<1$, vgl. § 6 dortselbst), es genügt also, wenn diese zu $\boldsymbol{R}(U)$ gehören; die $F(\varrho)$ aber entstanden durch Addieren, Subtrahieren, Multiplizieren aus den $E(U, a, b)$, $(-\infty<a,\ b<\infty$, vgl. §§ 4 u. 5 dortselbst), also sind nur diese von Interesse. $E(U, a, b)$ aber war Produkt von P. O. aus den Z. d. E. von $\frac{U+U^*}{2}$ und $\frac{U-U^*}{2i}$ [63]), und diese gehören nach unserem Satz 2 $\boldsymbol{R}(U)$ an, wenn $\frac{U+U^*}{2}$, $\frac{U-U^*}{2i}$ ihm angehören. Dieses ist klar: $\boldsymbol{R}(U)$ umfaßt U, U^*, also auch die letztgenannten O.

4. Nach diesen Vorbereitungen nehmen wir unseren eigentlichen Gegenstand wieder in Angriff. R, R^* sei wieder ein konj. Paar.

Satz 14. R, R^* sei normal. Wir bilden die zwei H. O. $S_1=\frac{R+R^*}{2}$, $S_2=\frac{R-R^*}{2i}$, und dann $\tilde{S}_1, \tilde{S}_2$ (vgl. Anfang Anm. [62])), die letzteren sind hypermax. Die O. $\overline{R}=\tilde{S}_1+i\tilde{S}_2$, $\overline{R}^*=\tilde{S}_1-i\tilde{S}_2$ (beide definiert im Durchschnitt der Definitionsbereiche von $\tilde{S}_1, \tilde{S}_2$) bilden ein konj. Paar, und zwar ein normales. $\overline{R}$, $\overline{R}^*$ sind Forts. von R, R^*.

Beweis: Wir bilden S_1, S_2 wie vorgeschrieben, der H. Charakter ist evident, und dann $\tilde{S}_1, \tilde{S}_2$. Gehöre A zu $\mathfrak{B}$. Wenn es mit R, R^* vert. ist, so ist es auch mit S_1, S_2 vert., also auch mit $\tilde{S}_1, \tilde{S}_2$ — hieraus folgert man $(R, R^*)' \subset (\tilde{S}_1)'$ und $(\tilde{S}_2)'$. Wegen $(R)'=(R^*)'$ (vgl. den Beweis von Satz 11) ist $(R)'=(R, R^*)'$, also $(R)' \subset (\tilde{S}_1)'$ und $(\tilde{S}_2)'$, $(R)'' \supset (\tilde{S}_1)''$ und $(\tilde{S}_2)''$. Mit $(R)''$ sind somit $(\tilde{S}_1)''$, $(\tilde{S}_2)''$ Abelsch, d. h. die konj. Paare $\tilde{S}_1, \tilde{S}_1$ und $\tilde{S}_2, \tilde{S}_2$ normal — also nach Satz 12 $\tilde{S}_1, \tilde{S}_2$ hypermax.

Da $\tilde{S}_1$, $\tilde{S}_2$ bzw. Forts. von S_1, S_2 sind, ist $\overline{R}, \overline{R}^*$ eine solche von R, R^*, daß $\overline{R}, \overline{R}^*$ ein konj. Paar ist, ist klar, da $\tilde{S}_1, \tilde{S}_2$ H. O. sind. Die Normalität erkennen wir so: Es ist $(\overline{R})'=(\overline{R}^*)'$, also beides $=(\overline{R}, \overline{R}^*)'$, und dieses offenbar $\supset(\tilde{S}_1, \tilde{S}_2)'$, welches seinerseits, wie wir schon wissen, $\supset(R)'$ ist. Also: $(\overline{R})' \supset (R)'$, $(\overline{R})'' \subset (R)''$, mit $(R)''$ ist daher auch $(\overline{R})''$ Abelsch, also $\overline{R}, \overline{R}^*$ normal.

Satz 15. $\overline{R}f$ (und $\overline{R}^*f$) ist dann und nur dann sinnvoll, wenn zwei f^*, f^{**} existieren, so daß für alle g mit sinnvollem Rg (und R^*g) $(f, Rg)=(f^{**}, g)$, $(f, R^*g)=(f^*, g)$ gilt. Und zwar ist dann $Rf=f^*$, $R^*f=f^{**}$.

---

[63]) $E(U, a, b)$ (a. a. O. heißt U A) ist als $P_{\mathfrak{M}(R,a)} \cdot P_{\mathfrak{N}(S,b)}$ $\left(R=\frac{U+U^*}{2}, S=\frac{U-U^*}{2i}\right)$ definiert, und es kommt nur auf die in Satz 8*, 9* dortselbst angegebenen Eigenschaften dieser P. O. an, welche den bzw. Z. d. E. offenbar zukommen.

Beweis: Die Notwendigkeit ist klar:

$$(f, Rg) = (f, \bar{\bar{R}}g) = (\bar{\bar{R}}^* f, g), \quad (f, R^* g) = (f, \bar{\bar{R}}^* g) = (\bar{\bar{R}} f, g).$$

Die Hinreichendheit erkennen wir so: $\frac{f^* + f^{**}}{2}$ und $\frac{f^* - f^{**}}{2i}$ sind nach Definition dem f durch S_1 bzw. S_2 zugeordnet (vgl. E., Definition 8, 9), also auch durch $\tilde{S}_1$, $\tilde{S}_2$. Da die letzteren hypermax. sind, ist $\tilde{S}_1 f$, $\tilde{S}_2 f$ sinnvoll, und zwar gleich $\frac{f^* + f^{**}}{2}$, $\frac{f^* - f^{**}}{2i}$. Also sind auch $\bar{\bar{R}} f$, $\bar{\bar{R}}^* f$ sinnvoll, und zwar gleich f^*, f^{**}.

Satz 16. Jede Forts. T, T^* von R, R^* (konj. Paar!, normal braucht sie nicht zu sein) hat $\bar{\bar{R}}$, $\bar{\bar{R}}^*$ zur Forts. $\bar{\bar{R}}$, $\bar{\bar{R}}^*$ ist (selbst als bloßes konj. Paar) max., und zwar die einzige max. Forts. von R, R^*.

Beweis: Die zwei letzten Behauptungen folgen aus der ersten, betrachten wir daher nur diese. Seien Tf, $T^* f$ sinnvoll, wenn wir sie gleich f^*, f^{**} setzen, so sind die Prämissen von Satz 15 erfüllt: bei sinnvollem Rg, $R^* g$ ist

$$(f, Rg) = (f, Tg) = (T^* f, g), \quad (f, R^* g) = (f, T^* g) = (Tf, g).$$

Also sind auch $\bar{\bar{R}} f$, $\bar{\bar{R}}^* f$ sinnvoll, und zwar gleich Tf, $T^* f$. D. h. $\bar{\bar{R}}$, $\bar{\bar{R}}^*$ ist Forts. von T, T^*. —

Bei normalen konj. Paaren von O. herrschen also viel einfachere Verhältnisse als bei den H. O. (vgl. E.): so ist die max. Forts. hier stets auf eine und nur eine Weise möglich. Wir werden uns daher im folgenden stets auf max. Paare R, R^* beschränken, diese sind nach Satz 16 durch $\bar{\bar{R}} = R$, $\bar{\bar{R}}^* = R^*$ charakterisiert. Für diese werden wir das Analogon der Hilbertschen Spektraldarstellung aufstellen.

V. Spektralform der normalen Operatoren.

1. Wir untersuchen, wie angekündigt, normale max. konj. Paare R, R^*.

Satz 17. Sei R, R^* ein normales max. konj. Paar. Es gibt dann eine Schar von P. O. $\Delta(z)$ (z durchläuft alle komplexen Zahlen) mit den folgenden Eigenschaften:

α) $\Delta(z') \cdot \Delta(z'') = \Delta(z'') \cdot \Delta(z') = \Delta(z)$

$$(\Re z = \mathrm{Min}(\Re z', \Re z''), \ \Im z = \mathrm{Min}(\Im z', \Im z'')).$$

β) $\Delta(z) f \to \Delta(z_0) f$ für alle f und $z \to z_0$, wenn dabei $\Re z \geqq \Re z_0$, $\Im z \geqq \Im z_0$ bleibt, und zwar ist die Konvergenz sowohl bei festem f, $\Re z$ und beliebigem $\Im z$, als auch bei festem f, $\Im z$ und beliebigem $\Re z$, gleichmäßig.

γ) $\Delta(z)f \to 0$ bzw. $\to f$ für alle f, wenn $\Re z \to -\infty$ oder $\Im z \to -\infty$ bzw. wenn $\Re z \to +\infty$ und $\Im z \to +\infty$.

(Auf Grund dieser Eigenschaften α) bis γ) heiße eine Schar von P.O. eine komplexe Z. d. E.)

δ) Rf (und $R^* f$) hat dann und nur dann Sinn, wenn

$$\iint |z|^2 d\,|\Delta(z)f|^2$$

endlich ist. ($\iint$ ist über die ganze komplexe Zahlenebene zu erstrecken. Da $\iint\limits_{\Re} d\,|\Delta(z)f|^2$ für jedes Rechteck $\Re \geqq 0$ ist[66]) und der Integrand $|z|^2$ stets $\geqq 0$ ist, ist das obige Integral seiner Natur nach entweder konvergent und endlich, oder eigentlich divergent und $+\infty$. Diese letztere Möglichkeit soll ausgeschlossen werden.)

ε) Im genannten Falle gilt für alle g

$$(Rf, g) = \iint z\,d(\Delta(z)f, g), \qquad (R^* f, g) = \iint \bar{z}\,d(\Delta(z)f, g).$$

(Die Integrale sind für solche f — und beliebige g — absolut konvergent[67]).)

(Auf Grund von δ) bis ε) sollen R, R^* und $\Delta(z)$ zusammengehörig heißen.)

---

[66]) Wenn das Rechteck $\Re$ die Ecken $a' + b'i$, $a' + b''i$, $a'' + b''i$, $a'' + b'i$ ($a' \leqq a''$, $b' \leqq b''$) hat, so ist

$$\begin{aligned}\iint\limits_{\Re} |\Delta(z)f|^2 = |\Delta(a'+b'i)f|^2 - |\Delta(a'+b''i)f|^2 \\ + |\Delta(a''+b''i)f|^2 - |\Delta(a''+b'i)f|^2\end{aligned}$$

und wegen $|Ef|^2 = (f, Ef)$ (E ein P.O.)

$$= |\{\Delta(a'+b'i) - \Delta(a'+b''i) + \Delta(a''+b''i) - \Delta(a''+b'i)\}f|^2 \geqq 0,$$

falls der O. $\{\ldots\}$ ein P.O. ist. In der Tat ist nach α)

$$\begin{aligned}\Delta(a'+b'i) - \Delta(a'+b''i) + \Delta(a''+b''i) - \Delta(a''+b'i) \\ = \Delta(a''+b''i)(1-\Delta(a'+b''i))(1-\Delta(a''+b'i)),\end{aligned}$$

womit auch dies in Evidenz gesetzt ist.

[67]) Alle diese Ausführungen stehen in weitestgehender Analogie zu denjenigen von E., Satz 36, mit denen man sie vergleichen möge. Der Vollständigkeit halber sei hier ein Beweis der absoluten Konvergenz von $\iint z\,d(\Delta(z)f, g)$ ($\iint \bar{z}\,d(\Delta(z)f, g)$ erledigt sich ebenso) gegeben.

Seien $\Re_1, \ldots, \Re_k$ irgendwelche Rechtecke (ohne gemeinsame innere Punkte), die zusammen einen endlichen Teil der Ebene ausfüllen; $\zeta_1, \ldots, \zeta_k$ bzw. Punkte derselben. Dann gilt ($h = 1, \ldots, k$, E_h sei der nach Anm. [66]) fürs Rechteck $\Re_h$ berechnete P.O.):

$$\begin{aligned}\Big|\iint\limits_{\Re_h} d(\Delta(z)f, g)\Big| = |(E_h f, g)| = |(E_h f, E_h g)| \leqq |E_h f| \cdot |E_h g| \\ = \sqrt{(E_h f, f)\cdot(E_h g, g)} = \sqrt{\iint\limits_{\Re_h} |d\Delta(z)f|^2 \cdot \iint\limits_{\Re_h} d\,|\Delta(z)g|^2},\end{aligned}$$

(Fortsetzung der Fußnote [67]) auf nächster Seite.)

ζ) Wenn die H.O. $\tilde{S}_1$, $\tilde{S}_2$ nach Satz 14 gebildet werden, so ist $\tilde{S}_1 f$ bzw. $\tilde{S}_2 f$ dann und nur dann sinnvoll, wenn

$$\iint (\Re z)^2 d\,|\Delta(z) f|^2 \quad \text{bzw.} \quad \iint (\Im z)^2 d\,|\Delta(z) f|^2$$

endlich ist. (Über die Natur dieser Integrale vgl. S. 411 und Anm. [64])).

η) Im genannten Falle gilt für alle g

$$(\tilde{S}_1 f, g) = \iint \Re z \cdot d(\Delta(z) f, g), \quad (\tilde{S}_2 f, g) = \iint \Im z \cdot d(\Delta(z) f, g).$$

(Die absolute Konvergenz der Integrale ist gesichert, vgl. ε) und Anm. [67])).

Beweis. Da $\tilde{S}_1$, $\tilde{S}_2$ hypermax. sind, gehören zu ihnen zwei Z. d. E. $E(\lambda)$, $F(\lambda)$ (vgl. E., Satz 36, sowie unsere Anm. [42])). Nach Definition ist $\tilde{S}_1 f$ bzw. $\tilde{S}_2 f$ dann sinnvoll, wenn $\int_{-\infty}^{\infty} \lambda^2 d\,|E(\lambda) f|^2$ bzw. $\int_{-\infty}^{\infty} \lambda^2 d\,|F(\lambda) f|^2$ endlich ist, und zwar ist dann für jedes g

$$(\tilde{S}_1 f, g) = \int_{-\infty}^{\infty} \lambda\, d(E(\lambda) f, g) \quad \text{bzw.} \quad (\tilde{S}_2 f, g) = \int_{-\infty}^{\infty} \lambda\, d(F(\lambda) f, g).$$

Die Menge der $E(\lambda)$ bzw. der $F(\lambda)$ heiße $\mathcal{E}$ bzw. $\mathcal{F}$. Jedes $E(\lambda)$ gehört zu $\mathcal{E}$, also zu $\mathcal{E}''$, also nach Satz 13 zu $(\tilde{S}_1)''$. Schon beim Beweise von Satz 14 sahen wir aber, daß $(\tilde{S}_1)'' \subset (R)''$ ist, also ist $\mathcal{E} \subset (R'')$. Ebenso ergibt sich $\mathcal{F} \subset (R)''$. Da $(R)''$ Abelsch ist, sind die Elemente von $\mathcal{E}$ mit denen von $\mathcal{F}$ vert.: d. h. jedes $E(a)$ mit jedem $F(b)$. Durch $\Delta(a + ib) = E(a) F(b) = F(b) E(a)$ definieren wir also eine Schar von P. O. $\Delta(z)$. Man überzeugt sich leicht davon, daß α) bis γ) gelten, d. h. daß dies eine komplexe Z. d. E. ist.

Ferner gehen die Integrale

$$\int_{-\infty}^{\infty} \lambda^2 d\,|E(\lambda) f|^2, \quad \int_{-\infty}^{\infty} \lambda^2 d\,|F(\lambda) f|^2, \quad \int_{-\infty}^{\infty} \lambda\, d(E(\lambda) f, g), \quad \int_{-\infty}^{\infty} \lambda\, d(F(\lambda) f, g)$$

---

und daraus folgt (vgl. die Abschätzungen in E., Beweis von Satz 36, sowie unsere Anm. [45]))

$$\left| \sum_{h=1}^{k} \zeta_h \iint_{\Re_h} d(\Delta(z) f, g) \right| \leq \sum_{h=1}^{k} |\zeta_h| \sqrt{\iint_{\Re_h} d\,|\Delta(z) f|^2 \cdot \iint_{\Re_h} d\,|\Delta(z) f|^2}$$

$$\leq \sqrt{\sum_{h=1}^{k} |\zeta_h|^2 \iint_{\Re_h} d\,|\Delta(z) f|^2} \cdot \sqrt{\sum_{h=1}^{k} \iint_{\Re_h} d\,|\Delta(z) g|^2}.$$

Da aber die Integrale

$$\iint |z|^2 d\,|\Delta(z) f|^2 \quad \text{und} \quad \iint d\,|\Delta(z) g|^2 = |g|^2$$

endlich sind, also absolut konvergieren (vgl. die Bemerkung zu δ)), folgt hieraus dasselbe für $\iint z\, d(\Delta(z) f, g)$.

in

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a^2 d|\Delta(a+ib)f|^2, \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b^2 d|\Delta(a+ib)f|^2,$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a\, d(\Delta(a+ib)f, g), \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b\, d(\Delta(a+ib)f, g)$$

über, also wenn $z = a + ib$ gesetzt wird, in diejenigen aus ζ) bis η). Also sind auch diese erfüllt.

Wegen $R = \bar{\tilde{R}}$, $R^* = \bar{\tilde{R}}^*$ hat Rf (und R^*f) dann und nur dann Sinn, wenn $\tilde{S}_1 f$, $\tilde{S}_2 f$ Sinn haben — d. i. wenn beide Integrale aus ζ) endlich sind. Nach dem, was dort über ihre wesentlich nicht-negative Natur gesagt wurde, bedeutet dies, daß ihre Summe endlich ist; und wegen $|z|^2 = (\Re z)^2 + (\Im z)^2$ ist das genau die Aussage von δ). ε) folgt ohne weiteres aus η).

Satz 18. Die Schar $\Delta(z)$ ist bereits durch die Bedingungen α) bis ε) eindeutig festgelegt: d. h. zu jedem max. konj. Paar R, R^* gehört eine und nur eine komplexe Z. d. E. $\Delta(z)$.

Beweis. Sei $\Delta(z)$ die beim Beweise von Satz 17 konstruierte Schar und $\Delta'(z)$ eine andere, die α) bis ε) auch erfüllt — es gilt $\Delta(z) = \Delta'(z)$ für alle z zu beweisen.

Zunächst folgt η) aus δ) zumindest für diejenigen f, für die Rf Sinn hat, d. h. $\tilde{S}_1 f$, $\tilde{S}_2 f$ beide Sinn haben — für diese f gilt es also auch mit $\Delta'(z)$. Das auf den Definitionsbereich von R eingeschränkte $\tilde{S}_1$ bzw. $\tilde{S}_2$ ist aber S_1 bzw. S_2 (beide haben denselben Definitionsbereich): also handelt es sich von den f, für die $S_1 f$, $S_2 f$ Sinn haben.

Da $\Delta'(a + ib)$ bei wachsendem b (und festem a!) im Sinne der P. O.-Relation $\leqq$ nie fällt (E. § III, Satz 14), so existiert für $b \to \infty$ ein starker Limes (E. § III, Satz 19): $\Delta'(a + ib) \to E'(a)$. Ebenso existiert bei festem b und $a \to \infty$ ein starker Limes: $\Delta'(a + ib) \to F'(b)$. Die $E'(a)$, $F'(b)$ sind P.O., die (ebenso wie $\Delta(a + ib)$, aus Stetigkeitsgründen) bei wachsendem a, b nie fallen; dabei ist für $a' \geqq a, b' \geqq b$ nach α)

$$\Delta'(a' + ib)\Delta'(a + ib') = \Delta'(a + ib')\Delta'(a' + ib) = \Delta'(a + ib),$$

also durch Grenzübergang ($a', b' \to +\infty$)

$$F'(b)E'(a) = E'(a)F'(b) = \Delta'(a + ib).$$

Aus

$$\Delta'(a+ib)f \to \begin{cases} E'(a)f & \text{für } b \to +\infty \\ 0 & \text{für } b \to -\infty \end{cases} \Bigg| \quad \Delta'(a+ib)f \to \begin{cases} F(b)f & \text{für } a \to +\infty \\ 0 & \text{für } a \to -\infty \end{cases} \Bigg|$$

folgert man sofort

$$\iint \Re z \cdot d(\varDelta'(z) f, g) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} a\, d(E'(a) F'(b) f, g) = \int_{-\infty}^{\infty} a\, d(E'(a) f, g),$$

$$\iint \Im z \cdot d(\varDelta'(z) f, g) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b\, d(E'(a) F'(b) f, g) = \int_{-\infty}^{\infty} b\, d(F'(b) f, g),$$

in dem Sinne, daß jede rechte Seite konvergiert, wenn es die entsprechende linke tut (die beiden Gleichungen gelten unabhängig voneinander).

Weiter schließt man mühelos aus α) bis γ), daß $E'(\lambda)$, $F'(\lambda)$ Z. d. E. sind, es mögen zu ihnen die H. O. T_1 bzw. T_2 gehören (vgl. E., Satz 36). Nach dem bisher Gesagten ist, wenn $T_1 f$ und $S_1 g$ beide Sinn haben,

$$(f, S_1 g) = \int_{-\infty}^{\infty} \lambda\, d(f, E(\lambda) g) = \int_{-\infty}^{\infty} \lambda\, d(E(\lambda) f, g) = (T_1 f, g),$$

d. h. f ist Erweiterungselement von S_1, und zwar wird ihm durch S_1 $T_1 f$ zugeordnet — also auch durch $\tilde{S}_1$. Aber $\tilde{S}_1$ ist hypermax.: also ist $\tilde{S}_1 f$ sinnvoll und gleich $T_1 f$. D. h.: $\tilde{S}_1$ ist Forts. von T_1 und, da auch T_1 hypermax. ist, $\tilde{S}_1 = T_1$. Ebenso zeigt man $\tilde{S}_2 = T_2$. Da nun $\tilde{S}_1$, $\tilde{S}_2$ die bzw. Z.d.E. $E(\lambda)$, $F(\lambda)$ haben, muß $E'(\lambda) = E(\lambda)$, $F'(\lambda) = F(\lambda)$ sein. Hieraus folgt aber

$$\varDelta'(a + ib) = E'(a) F'(b) = E(a) F(b) = \varDelta(a + ib),$$

d. h. $\varDelta'(z) = \varDelta(z)$, was zu beweisen war.

Satz 19. Zu jeder komplexen Z. d. E., d. h. jeder α) bis γ) erfüllenden Schar $\varDelta(z)$ gehört ein und nur ein konj. Paar R, R^* (nach δ) bis ε)), und dieses ist max. und normal.

Beweis. Wenn es ein solches Paar gibt, gibt es nur eines: denn δ) legt seinen Definitionsbereich fest, und ε) die (Rf, g), $(R^* f, g)$, d. h. die Rf, $R^* f$ selbst. Bezüglich der Existenz ist noch zu sagen: wenn wir irgendein O.-Paar R, R^* nach δ), ε) haben (und der Definitionsbereich überall dicht ist), so bilden diese nach δ), ε) von selbst ein konj. Paar — nur die Max. und Normalität ist dann noch zu prüfen.

Die Existenz der R, R^* nach δ), ε) steht fest, sobald wir dieses wissen: wenn $\iint |z|^2 d|\varDelta(z) f|^2$ endlich ist, so gibt es zwei f^*, f^{**}, so daß für alle g

$$\iint z\, d(\varDelta(z) f, g) = (f^*, g), \qquad \iint \bar{z}\, d(\varDelta(z) f, g) = (f^{**}, g)$$

gilt (dieses ist dann Rf bzw. $R^* f$). Wir nennen die Integrale auf den linken Seiten vorübergehend $L^*(g)$, $L^{**}(g)$ (f sei fest gedacht). Da offenbar

$$L^*(a_1 g_1 + \ldots + a_n g_n) = \bar{a}_1 L^*(g_1) + \ldots + \bar{a}_n L^*(g_n),$$
$$L^{**}(a_1 g_1 + \ldots + a_n g_n) = \bar{a}_1 L^{**}(g_1) + \ldots + \bar{a}_n L^{**}(g_n)$$

gilt, sind wir nach dem Satz von F. Riesz (vgl. E., Anm. [52])) am Ziele, wenn wir eine Konstante c mit $|L^*(g)| \leqq c \cdot |g|$, $|L^{**}(g)| \leqq c \cdot |g|$ finden. Aber aus der Abschätzung von Anm. [67]) folgt

$$|L^*(g)| \leqq \sqrt{\iint |z^2| d|\Delta(z)f|^2} \sqrt{\iint d|\Delta(z)g|^2} = \sqrt{\iint |z|^2 d|\Delta(z)f|^2} \cdot |g|,$$

und ebenso

$$|L^{**}(g)| \leqq \sqrt{\iint |z|^2 d|\Delta(z)f|^2} \cdot |g|,$$

womit alles Gewünschte erreicht ist.

Der Definitionsbereich, d. h. die f mit endlichem $\iint |z|^2 d|\Delta(z)f|^2$, ist überall dicht. Dies zeigt man so: Wie beim Beweise von Satz 18 zu $\Delta'(z)$ die $E'(\lambda)$, $F'(\lambda)$ gebildet wurden, bilden wir jetzt zu $\Delta(z)$ $E(\lambda)$, $F(\lambda)$ zwei Z. d. E. $\iint (z)^2 d|\Delta(z)f|^2$ ist endlich, wenn $\iint (\Re z)^2 d|\Delta(z)f|^2$, $\iint (\Im z)^2 d|\Delta(z)f|^2$ es sind, d. h. wenn $\int_{-\infty}^{\infty} \lambda^2 d|E(\lambda)f|^2$, $\int_{-\infty}^{\infty} \lambda^2 d|F(\lambda)f|^2$ (vgl. dort) es sind. Das trifft gewiß für alle f mit $E(\lambda)f$ und $F(\lambda)f = \begin{cases} f & \text{für } \lambda \geqq c \\ 0 & \text{für } \lambda \leqq -c \end{cases}$ (c beliebig, aber fest) zu[68]) — also für jedes $f = (E(c) - E(-c))(F(c) - F(-c))g$. Da dieses für $c \to \infty$ gegen g konvergiert und g beliebig ist, ist die betrachtete Menge wirklich überall dicht.

Es bleibt noch zu zeigen, daß R, R^* max. und normal ist. Zunächst die Normalität. Sei $\mathfrak{D}$ die Menge aller $\Delta(z)$, da diese H. und vert. sind, ist $\mathfrak{D}$ Abelsch, also auch $\mathfrak{D}''$. Daher genügt es, $(R)'' \subset \mathfrak{D}''$ zu beweisen, also erst recht $(R)' \supset \mathfrak{D}'$. Dies ist erwiesen, wenn wir zeigen: wenn A mit allen $\Delta(z)$ vert. ist, so ist es auch mit R vert.

Sei also Rf sinnvoll, wir behaupten: $R(Af)$ ist sinnvoll und gleich $A(Rf)$. Wir müssen also erstens aus der Endlichkeit von $\iint |z|^2 d|\Delta(z)f|^2$ auf diejenige von $\iint |z|^2 d|\Delta(z)Af|^2$ schließen und dann etwa $(R(Af), g) = (A(Rf), g)$ (für alle g!), d. h. $(Af, R^*g) = (Rf, A^*g)$, also $\iint z\, d(Af, \Delta(z)g) = \iint z\, d(\Delta(z)f, A^*g)$ beweisen.

A ist beschränkt, d. h. stets $|Af| \leqq c \cdot |f|$, hieraus folgt für jedes Rechteck $\Re$ (wenn $E_\Re$ der nach Anm. [66]) dazu konstruierte P. O. ist)

$$\iint_\Re d|\Delta(z)Af|^2 = |E_\Re Af|^2 = |AE_\Re f|^2 \leqq c^2 |E_\Re f|^2 = c^2 \iint_\Re d|\Delta(z)f|^2;$$

wegen $\lambda^2 \geqq 0$ ist also $\iint \lambda^2 d|\Delta(z)Af|^2 \leqq c^2 \iint \lambda^2 d|\Delta(z)f|^2$, womit die erste Behauptung bewiesen ist. Die zweite dagegen ist nach

$$(Af, \Delta(z)g) = (\Delta(z)Af, g) = (A\Delta(z)f, g) = (\Delta(z)f, A^*g)$$

trivial.

---

[68]) Denn dann können die Integrale $\int_{-\infty}^{\infty}$ durch $\int_{-c}^{c}$ ersetzt werden, und dann ist der Integrand λ^2 beschränkt.

Betrachten wir nun die Max. Diese besagt $\bar{\bar{R}} = R$, $\bar{\bar{R}}^* = R^*$, d. h.: da der Durchschnitt der Definitionsbereiche von $\tilde{S}_1, \tilde{S}_2$ (vgl. Satz 14) der Definitionsbereich von $\bar{\bar{R}}, \bar{\bar{R}}^*$ ist, ist zu verlangen, daß R (und damit R^*) in ihm überall sinnvoll ist. Wir ziehen wieder die Z. d. E. $E(\lambda)$, $F(\lambda)$ von vorhin heran, zu ihnen mögen die (hypermax.) H. O. T_1, T_2 gehören. Offenbar ist T_1 Forts. von S_1, also (da es abg. lin. ist) auch von $\tilde{S}_1$ und, da dieses hypermax. ist, $T_1 = S_1$ — ebenso zeigt man $T_2 = S_2$. Daher sind im Durchschnitt der Definitionsbereiche von $\tilde{S}_1 f$, $\tilde{S}_2 f$ beide $T_1 f$, $T_2 f$ sinnvoll, d. h. die Integrale

$$\int_{-\infty}^{\infty} \lambda^2 d\,|E(\lambda) f|^2 = \iint (\mathfrak{R}\, z)^2 d\,|\Delta(z) f|^2,$$

$$\int_{-\infty}^{\infty} \lambda^2 d\,|F(\lambda) f|^2 = \iint |\mathfrak{J}\, z|^2 d\,|\Delta(z) f|^2$$

endlich — also auch ihre Summe $\iint |z|^2 d\,|\Delta(z) f|^2$. Somit ist Rf sinnvoll, also alles bewiesen.

Die Sätze 17 bis 19 zeigen, daß die max. normalen konj. Paare von O. den komplexen Z. d. E. in ebenso ein-eindeutiger Weise entsprechen, wie die hypermax. H. O. den reellen Z. d. E. (nach E., Satz 36). Wegen der allgemeinen Eindeutigkeit der max. Forts. (und dem Fehlen eines Analogons der Hypermax.) sind die Verhältnisse sogar etwas einfacher.

2. Wir stellen noch einige einfache Eigenschaften der max. normalen konj. Paare fest.

Satz 20. R, R^* sei ein max. normales konj. Paar, $\Delta(z)$ seine komplexe Z. d. E., $\mathfrak{D}$ die Menge aller $\Delta(z)$. Dann ist $(R)' = \mathfrak{D}'$, $(R)'' = \mathfrak{D}''$, $(R)''' = \mathfrak{D}'''$, ...

Beweis. Es genügt offenbar, die erste Gleichung zu beweisen, und $(R)' \supset \mathfrak{D}'$ sahen wir schon beim vorigen Beweise — es bleibt also noch $(R)' \subset \mathfrak{D}'$ zu zeigen. Schon beim Beweise von Satz 14 kam $(R)' \subset (\tilde{S}_1)'$ und $(\tilde{S}_2)'$, d. h. $(R)' \subset (\tilde{S}_1, \tilde{S}_2)'$ vor, es genügt also, $(\tilde{S}_1, \tilde{S}_2)' \subset \mathfrak{D}'$ einzusehen. Die Z. d. E. von $\tilde{S}_1$ bzw. $\tilde{S}_2$ sei $E(\lambda)$ bzw. $F(\lambda)$, die Menge aller $E(\lambda)$ bzw. $F(\lambda)$ $\mathcal{E}$ bzw. $\mathcal{F}$; nach Satz 13 ist $(\tilde{S}_1)' = \mathcal{E}'$, $(\tilde{S}_2)' = \mathcal{F}'$, also $(\tilde{S}_1, \tilde{S}_2)' = (\mathcal{E}, \mathcal{F})'$.

Wenn nun A (aus $\mathfrak{B}$) mit allen Elementen von $\mathcal{E}$ und $\mathcal{F}$ vert. ist, so ist es auch mit jedem $\Delta(a + ib) = E(a) F(b)$ vert., d. h. mit jedem Element von $\mathfrak{D}$. Somit ist $(\mathcal{E}, \mathcal{F})' \subset \mathfrak{D}'$. Damit ist alles bewiesen.

Satz 21. R, R^* sei ein normales konj. Paar, dann gilt allgemein $|Rf| = |R^* f|$. Wenn R, R^* auch max. ist, so ist sowohl R als auch R^* ein abg. lin. O.

Beweis. Da R, R^* allenfalls eine max. Forts. besitzt (Satz 16), genügt es, die erste Behauptung für max. R, R^* zu beweisen. Sei also $\Delta(z)$

seine komplexe Z. d. E. Wenn Rf (und $R^* f$) Sinn hat, so gilt

$$\begin{aligned}(Rf, \Delta(a+ib)g) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a'+ib')\,d(\Delta(a'+ib')f, \Delta(a+ib)g)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a'+ib')\,d(\Delta(a+ib)\Delta(a'+ib')f, g)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a'+ib')\,d(\Delta(\mathrm{Min}(a,a')+i\,\mathrm{Min}(b,b'))f, g)\\ &= \int_{-\infty}^{a}\int_{-\infty}^{b}(a'+ib')\,d(\Delta(a+ib)f, g),\end{aligned}$$

also insbesondere

$$\begin{aligned}(\Delta(a+ib)f, Rf) &= \int_{-\infty}^{a}\int_{-\infty}^{b}(a'-ib')\,d(f, \Delta(a+ib)f)\\ &= \int_{-\infty}^{a}\int_{-\infty}^{b}(a'-ib')\,d\,|\Delta(a+ib)f|^2.\end{aligned}$$

Daraus folgt weiter

$$\begin{aligned}|Rf|^2 = (Rf, Rf) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a+ib)\,d(\Delta(a+ib)f, Rf)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a+ib)\,d\left[\int_{-\infty}^{a}\int_{-\infty}^{b}(a'-ib')\,d\,|\Delta(a'+ib')f|^2\right]^{69)}\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a+ib)(a-ib)\,d\,|\Delta(a+ib)f|^2\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a^2+b^2)\,d\,|\Delta(a+ib)f|^2.\end{aligned}$$

Ebenso zeigt man

$$|R^* f|^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a^2+b^2)\,d\,|\Delta(a+ib)f|^2.$$

Damit ist $|Rf| = |R^* f|$ in der Tat bewiesen.

Gehen wir nun zur zweiten Behauptung über. Daß R wie R^* lin. ist, folgt z. B. aus Satz 15 (wegen der Max. ist ja $R = \tilde{R}$, $R^* = \tilde{R}^*$). Wir zeigen jetzt die Abg. von R — die von R^* wird ebenso bewiesen. Sei $f_n \to f$, alle Rf_n (und daher $R^* f_n$) sinnvoll, $Rf_n \to f^*$. Also erfüllen die Rf_n die Cauchysche Konvergenzbedingung ($|R(f_m - f_n)| \to 0$ für $m, n \to \infty$), die $R^* f_m$ daher auch (weil $|R(f_m - f_n)| = |R^*(f_m - f_n)|$ ist) — also existiert ein f^{**} mit $R^* f_n \to f^{**}$. Hieraus folgt sofort, daß f, f^*, f^{**} die Prämissen von Satz 15 erfüllen, d. h. daß Rf sinnvoll und gleich f^* ist. Also ist R abg., wie behauptet wurde.

[69]) Zu dieser Integral-Umformung vgl. E., Anm. [55]).

Anhang I.

Der Vollständigkeit halber soll hier die bekannte (Hilbertsche) Kennzeichnung der Z. d. E. der beschr. H. O. gegeben werden. Sei also R hypermax., $E(\lambda)$ eine Z. d. E.

Nach E. § II, Satz 12γ ist für die Beschr. von R $|(Rf, f)| \leqq c \cdot |f|^2$ (c fest!) charakteristisch, und dabei ist

$$(Rf, f) = \int_{-\infty}^{\infty} \lambda \, d(E(\lambda) f, f) = \int_{-\infty}^{\infty} \lambda \, d |E(\lambda) f|^2 .$$

Wenn $E(\lambda) = \left\{ \begin{matrix} 1 & \text{für } \lambda \geqq c \\ 0 & \text{für } \lambda \leqq -c \end{matrix} \right|$ ist, so können wir $\int_{-\infty}^{\infty}$ durch $\int_{-c}^{c}$ ersetzen; dann bleibt der Integrand absolut $\leqq c$, also das Integral ($|E(\lambda)f|^2$ nimmt nie ab!) absolut $\leqq c \int_{-\infty}^{\infty} d|E(\lambda)f|^2 = c \cdot |f|^2$ — d. h. R ist beschr.

Ist umgekehrt $E(\lambda) \neq 0$ bzw. $\neq 1$ für beliebig kleine bzw. große λ, also für absolut beliebig große λ nicht konstant (wegen $E(\lambda) f \to 0$ bzw. f für $\lambda \to -\infty$ bzw. $+\infty$ kommt ja kein anderer konstanter Wert für $E(\lambda)$ als 0 bzw. 1 in Frage!), so kommt $E(\lambda) \neq E(\mu)$ und $\lambda < \mu < -D$ oder $D < \lambda < \mu$ für jedes D vor. Somit ist $E(\mu) - E(\lambda)$ ein P. O. und $\neq 0$; sei $f \neq 0$ ein Punkt seiner abg. lin. M., d. h. $(E(\mu) - E(\lambda)) f = f$. Dann ist $E(\lambda') f = \left\{ \begin{matrix} f & \text{für } \lambda' \geqq \mu \\ 0 & \text{für } \lambda' \leqq \lambda \end{matrix} \right|$.[70]) Also haben wir

$$|(Rf, f)| = \left| \int_{-\infty}^{\infty} \lambda' d |E(\lambda') f|^2 \right| = \left| \int_{\lambda}^{\mu} \lambda' d |E(\lambda') f|^2 \right| \geqq D \int_{\lambda}^{\mu} d |E(\lambda) f|^2$$

$$= D \int_{-\infty}^{\infty} d |E(\lambda) f|^2 = D \cdot |f|^2 .$$

Für $D > c$ ist dies $> c \cdot |f|^2$, d. h. R unbeschr.

Es ist ein leichtes, analog die Beschr. für max. normale konj. Paare R, R^* an der komplexen Z. d. E. $\Delta(z)$ abzulesen. Sie ist durch $|Rf| \leqq c \cdot |f|$ (c fest!) definiert, man legt vorteilhaft die folgende Relation zugrunde (nach Satz 21):

$$|Rf|^2 = \iint |z|^2 d |\Delta(z) f|^2 .$$

---

[70]) Es ist ja

$$|E(\mu) f|^2 - |E(\lambda) f|^2 = |(E(\mu) - E(\lambda)) f|^2 = |f|^2 ,$$

$$|E(\mu) f|^2 \leqq |f|^2 , \quad |E(\lambda) f|^2 \geqq 0$$

(vgl. E. § III), also $|E(\mu) f| = |f|$, $|E(\lambda) f| = 0$. Erst recht ist daher für $\lambda' \geqq \mu$ bzw. $\leqq \lambda$ $|E(\lambda') f| = |f|$ bzw. 0 — d. h. $E(\lambda') f = f$ bzw. 0 (vgl. a. a. O.), wie behauptet.

Anhang II.

1. Es soll eine Anwendung der Resultate der §§ IV—V gemacht werden, insbesondere auf Laurentsche Matrizen und deren Verallgemeinerungen.

Sei $\mathfrak{P}$ eine Teilmenge von $\mathfrak{B}$, für welche $\mathfrak{P}'$ Abelsch ist. Wenn R, R^* ein konj. Paar ist, so sagen wir, daß es von der Klasse $\mathfrak{P}$ ist, falls $(R)' \supset \mathfrak{P}$ gilt — d. h. wenn R mit allen A, A^* vert. ist, falls A ganz $\mathfrak{P}$ durchläuft.

Da dann $(R)'' \subset \mathfrak{P}'$ ist, ist $(R)''$ Abelsch, d. h. R, R^* normal; nach Satz 16 hat es somit eine einzige max. Forts., das ebenfalls normale $\overline{R}, \overline{R}^*$. Beim Beweise von Satz 14 sahen wir $(\overline{R})' \supset (R)'$, also $\supset \mathfrak{P}$ — d. h. auch $\overline{R}$ ist von der Klasse $\mathfrak{P}$. Somit genügt es, max. normale R, R^* zu betrachten. Als solches werde R, R^* vorausgesetzt, seine komplexe Z. d. E. sei $\Delta(z)$, $\mathfrak{D}$ die Menge aller $\Delta(z)$. Wegen $(R)' = \mathfrak{D}'$ (Satz 20) ist R, R^* dann und nur dann von der Klasse $\mathfrak{P}$, wenn $\mathfrak{D}' \supset \mathfrak{P}$ ist. (Wegen $(\Delta(z)) \subset \mathfrak{D}$, $(\Delta(z))' \supset \mathfrak{D}' \supset \mathfrak{P}$ gehört also dann jedes $\Delta(z)$ zur Klasse $\mathfrak{P}$ — da $\mathfrak{D}'$ der Durchschnitt aller $(\Delta(z))'$ ist, ist dies auch hinreichend.) Aus $\mathfrak{D}' \supset \mathfrak{P}$ folgt $\mathfrak{P}' \supset \mathfrak{D}'' \supset \mathfrak{D}$, aus $\mathfrak{P}' \supset \mathfrak{D}$ $\mathfrak{D}' \supset \mathfrak{P}'' \supset \mathfrak{P}$, also ist auch $\mathfrak{D} \subset \mathfrak{P}'$, d. h, die Zugehörigkeit aller $\Delta(z)$ zu $\mathfrak{P}'$, notwendig und hinreichend. —

Wir stellen nun ein $\mathfrak{P}$ von der oben genannten Art auf die folgende Weise her. Sei $\mathfrak{G}$ eine abzählbar-unendliche Abelsche Gruppe (die Methode ist auch auf kontinuierliche Gruppen $\mathfrak{G}$ verallgemeinerbar, was hier nicht erörtert werde), $\alpha, \beta, \ldots$ ihre Elemente. Sei $\varphi_\alpha, \varphi_\beta, \ldots$ ein vollst. norm. orth. System (in $\mathfrak{H}$), das wir nicht mit den Zahlen $1, 2, \ldots$, sondern mit den Elementen $\alpha, \beta, \ldots$ von $\mathfrak{G}$ indizieren.

Man sieht sofort, daß der O. U_α, der durch

$$U_\alpha\Big(\sum_{\beta \text{ in } \mathfrak{G}} x_\beta \varphi_\beta\Big) = \sum_{\beta \text{ in } \mathfrak{G}} x_{\alpha\beta} \varphi_\beta \qquad \Big(\sum_{\beta \text{ in } \mathfrak{G}} |x_\beta|^2 \text{ endlich!}\Big)$$

definiert wird, ein unitärer ist. Dabei gilt $U_\alpha U_\beta = U_{\alpha\beta}$, also insbesondere $U_\alpha^* = U_\alpha^{-1} = U_{\alpha^{-1}}$. Sei nun $\mathfrak{P}$ die Menge aller U_α. Wann gehört ein A von $\mathfrak{B}$ zu $\mathfrak{P}'$? A muß mit jedem U_α von $\mathfrak{P}$ (die $U_\alpha^* = U_{\alpha^{-1}}$ liefern nichts Neues!) vert. sein, $AU_\alpha = U_\alpha A$. Dies bedeutet: $(AU_\alpha f, g) = (U_\alpha A f, g)$ für alle f, g — da aber beide Seiten linear und stetig in f und in g sind, genügt es, dies für ein vollst. norm. orth. System zu fordern. Z. B. für alle $f = \varphi_\beta$, $g = \varphi_\gamma$, wegen

$$(AU_\alpha \varphi_\beta, \varphi_\gamma) = (A\varphi_{\alpha^{-1}\beta}, \varphi_\gamma),$$

$$(U_\alpha A\varphi_\beta, \varphi_\gamma) = (A\varphi_\beta, U_\alpha^* \varphi_\gamma) = (A\varphi_\beta, U_{\alpha^{-1}}\varphi_\gamma) = (A\varphi_\beta, \varphi_{\alpha\gamma})$$

bedeutet das $(A\varphi_{\alpha^{-1}\beta}, \varphi_\gamma) = (A\varphi_\beta, \varphi_{\alpha\gamma})$ — d. h. daß $(A\varphi_\beta, \varphi_\gamma)$ nur von $\beta^{-1}\gamma$ abhängt.

Wenn also A, B zu $\mathfrak{P}$ gehören, so ist (wir setzen $(A\varphi_\varepsilon, \varphi_\eta) = a(\varepsilon^{-1}\eta)$, $(B\varphi_\varepsilon, \varphi_\eta) = b(\varepsilon^{-1}\eta)$)

$$(AB\varphi_\beta, \varphi_\gamma) = (B\varphi_\beta, A^*\varphi_\gamma) = \sum_{\delta \text{ in } \mathfrak{G}} (B\varphi_\beta, \varphi_\delta)(\varphi_\delta, A^*\varphi_\gamma)$$

$$= \sum_{\delta \text{ in } \mathfrak{G}} (B\varphi_\beta, \varphi_\delta)(A\varphi_\delta, \varphi_\gamma) = \sum_{\delta \text{ in } \mathfrak{G}} a(\delta^{-1}\gamma)\, b(\beta^{-1}\delta) = \sum_{\varepsilon, \eta \text{ in } \mathfrak{G},\ \varepsilon\eta = \beta^{-1}\gamma} a(\varepsilon)\, b(\eta),$$

und genau so kommt

$$(BA\varphi_\beta, \varphi_\gamma) = \sum_{\varepsilon, \eta \text{ in } \mathfrak{G},\ \varepsilon\eta = \beta^{-1}\gamma} a(\varepsilon)\, b(\eta)$$

heraus. Daher gilt $(ABf, g) = (BAf, g)$ zumindest im vollst. norm. orth. System der φ_α, also, da beide Seiten linear und stetig in f und in g sind, für alle f, g. Somit ist $AB = BA$, A, B sind vert., und damit ist $\mathfrak{P}'$ (welches ja mit A auch A^* enthält) als Abelsch erkannt.

2. Auf die soeben konstruierte Menge $\mathfrak{P}$ wollen wir nun das am Anfang des § 1 Gesagte anwenden, d. h. die R, R^* dieser Klasse untersuchen.

Sei a_α eine Zahlenfolge (gleichfalls mit Elementen α von $\mathfrak{G}$ indiziert), wobei nur die Endlichkeit von $\sum_{\alpha \text{ in } \mathfrak{G}} |a_\alpha|^2$ vorausgesetzt werde. Wir definieren zwei O. R, R^* für die φ_α, und nur diese, folgendermaßen:

$$R\varphi_\alpha = \sum_{\beta \text{ in } \mathfrak{G}} a_{\alpha^{-1}\beta}\varphi_\beta, \qquad R^*\varphi_\alpha = \sum_{\beta \text{ in } \mathfrak{G}} \overline{a_{\alpha\beta^{-1}}}\varphi_\beta$$

($\sum_{\beta \text{ in } \mathfrak{G}} |a_{\alpha^{-1}\beta}|^2 = \sum_{\beta \text{ in } \mathfrak{G}} |a_{\alpha\beta^{-1}}|^2 = \sum_{\beta \text{ in } \mathfrak{G}} |a_\beta|^2$ ist endlich). Man sieht sofort, daß R, R^* im Sinne von Definition 5 ein konj. Paar bilden; ferner, daß R mit allen U_α (also auch den $U_\alpha^* = U_{\alpha^{-1}}$) vert. ist — d. h. $(R)' \supset \mathfrak{P}$ gilt. Also gehört R, R^* zur Klasse $\mathfrak{P}$; es ist normal, wir können daher $\overline{\overline{R}}, \overline{\overline{R}}^*$ bilden. Dieses ist max. normal und gehört auch zur Klasse $\mathfrak{P}$.

Nun bestimmen wir, auf Grund von Satz 15, den Definitionsbereich und den Wertverlauf von $\overline{\overline{R}}$ und $\overline{\overline{R}}^*$. Danach sind $\overline{\overline{R}}f$, $\overline{\overline{R}}^*f$ dann und nur dann sinnvoll, wenn zwei f^*, f^{**} existieren, so daß (soweit Rg, R^*g Sinn haben)

$$(f, Rg) = (f^{**}, g), \qquad (f, R^*g) = (f^*, g)$$

gilt. Da $g = \varphi_\alpha$ sein muß, bedeutet dies, wenn wir noch $f = \sum_{\beta \text{ in } \mathfrak{G}} x_\beta \varphi_\beta$ setzen ($\sum_{\beta \text{ in } \mathfrak{G}} |x_\beta|^2$ endlich):

$$\sum_{\beta \text{ in } \mathfrak{G}} x_\beta \overline{a_{\alpha^{-1}\beta}} = (f^{**}, \varphi_\alpha), \qquad \sum_{\beta \text{ in } \mathfrak{G}} x_\beta a_{\alpha\beta^{-1}} = (f^*, \varphi_\alpha).$$

Wir setzen

$$y_\alpha = \sum_{\beta \text{ in } \mathfrak{G}} a_{\alpha\beta^{-1}} x_\beta, \quad z_\alpha = \sum_{\beta \text{ in } \mathfrak{G}} \overline{a_{\alpha^{-1}\beta}} x_\beta,$$

dann muß also $(f^*, \varphi_\alpha) = y_\alpha$, $(f^{**}, \varphi_\alpha) = z_\alpha$ sein — nach E. § I, Satz 5, 7 ist dies dann und nur dann möglich, wenn $\sum_{\alpha \text{ in } \mathfrak{G}} |y_\alpha|^2$, $\sum_{\alpha \text{ in } \mathfrak{G}} |z_\alpha|^2$ endlich sind, und zwar ist dann $f^* = \sum_{\alpha \text{ in } \mathfrak{G}} y_\alpha \varphi_\alpha$, $f^{**} = \sum_{\alpha \text{ in } \mathfrak{G}} z_\alpha \varphi_\alpha$. Unser Resultat ist also:

$\overline{\overline{R}} f$, $\overline{\overline{R}}^* f$ haben für $f = \sum_{\alpha \text{ in } \mathfrak{G}} x_\alpha \varphi_\alpha$ ($\sum_{\alpha \text{ in } \mathfrak{G}} |x_\alpha|^2$ endlich) dann und nur dann Sinn, wenn für die soeben definierten y_α, z_α die Summen $\sum_{\alpha \text{ in } \mathfrak{G}} |y_\alpha|^2$, $\sum_{\alpha \text{ in } \mathfrak{G}} |z_\alpha|^2$ endlich sind, und zwar ist dann $Rf = \sum_{\alpha \text{ in } \mathfrak{G}} y_\alpha \varphi_\alpha$, $R^* f = \sum_{\alpha \text{ in } \mathfrak{G}} z_\alpha \varphi_\alpha$.

Andererseits besitzt aber $\overline{\overline{R}}$, $\overline{\overline{R}}^*$ nach Satz 17, 18 genau eine Z. d. E. $\Delta(z)$; wie in § 1 festgestellt wurde, gehören die $\Delta(z)$ zu P. In unserem Falle bedeutet dies, wie wir ebendort sahen, daß $(\Delta(z)\varphi_\alpha, \varphi_\beta)$ nur von $\alpha^{-1}\beta$ abhängt: dieses ist das allgemeine Element der Matrix des P. O. $\Delta(z)$ im vollst. norm. orth. System $\varphi_\alpha, \varphi_\beta, \ldots$. —

Wählen wir für $\mathfrak{G}$ z. B. die unendliche zyklische Gruppe (z. B. alle ganzen Zahlen $0, \pm 1, \pm 2, \ldots$, mit der Addition als Kompositionsregel), so gewinnen wir eine allgemeine Eigenwerttheorie der Laurentschen Matrizen[71]) (auch der nicht-reellen und unbeschränkten). Andere Wahlen von $\mathfrak{G}$ führen zu analogen, aber allgemeineren Matrizenklassen, die wir nun gleicherweise beherrschen.

Anhang III.

1. Hier soll der Vertauschbarkeitsbegriff für unbeschr. O. etwas näher analysiert werden. Wir hatten die Normalität von R, d. i. die Vert. von R, R^*, durch den Abelschen Charakter von $(R)''$ definiert — also wenn wir die H. O. $S_1 = \frac{R+R^*}{2}$, $S_2 = \frac{R-R^*}{2i}$ bilden, deren Vert. durch den Abelschen Charakter von $(S_1, S_2)''$.[72]) Es liegt aber nahe, zu diesem Zwecke die übliche Definition der Matrizenvert. heranzuziehen zu versuchen.

Seien z. B. R, S zwei H. O., $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System, in welchem beide eine Matrix besitzen: $\{a_{\mu\nu}\}$ bzw. $\{b_{\mu\nu}\}$.[73]) Da die Zeilen- und Spalten-Absolutwertquadratsummen $\sum_{\mu=1}^{\infty} |a_{\mu\nu}|^2 = \sum_{\mu=1}^{\infty} |a_{\nu\mu}|^2$, $\sum_{\mu=1}^{\infty} |b_{\mu\nu}|^2 = \sum_{\mu=1}^{\infty} |b_{\nu\mu}|^2$ alle endlich sind, konvergieren auch die Reihen

---

[71]) Diese und ähnliche Klassen von Matrizen hat Toeplitz mit funktionentheoretischen Fragen in Beziehung gebracht. Vgl. auch a. a. O. Anm. [24]).

[72]) Beim Beweise von Satz 14 sahen wir $(R)' = (R, R^*)'$, also $= (S_1, S_2)'$, $(R)'' = (S_1, S_2)''$.

[73]) Vgl. E., Anhang III, ferner die demnächst im Journal f. Math. erscheinende Arbeit des Verfassers „Zur Theorie der unbeschränkten Matrizen".

$\sum_{\varrho=1}^{\infty} a_{\mu\varrho} b_{\varrho\nu}$, $\sum_{\varrho=1}^{\infty} b_{\mu\varrho} a_{\varrho\nu}$ absolut, und man könnte versuchen, die Vert. durch

$$\sum_{\varrho=1}^{\infty} a_{\mu\varrho} b_{\varrho\nu} = \sum_{\varrho=1}^{\infty} b_{\mu\varrho} a_{\varrho\nu} \qquad (\mu, \nu = 1, 2, \ldots)$$

zu definieren.

Wegen $a_{\mu\nu} = (R\varphi_\mu, \varphi_\nu)$, $b_{\mu\nu} = (S\varphi_\mu, \varphi_\nu)$ (vgl. z. B. E., Anhang III) ist aber

$$\sum_{\varrho=1}^{\infty} a_{\mu\varrho} b_{\varrho\nu} = \sum_{\varrho=1}^{\infty} (R\varphi_\mu, \varphi_\varrho)(S\varphi_\varrho, \varphi_\nu) = \sum_{\varrho=1}^{\infty} (R\varphi_\mu, \varphi_\varrho)(\varphi_\varrho, S\varphi_\nu) = (R\varphi_\mu, S\varphi_\nu),$$

$$\sum_{\varrho=1}^{\infty} b_{\mu\varrho} a_{\varrho\nu} = \sum_{\varrho=1}^{\infty} (S\varphi_\mu, \varphi_\varrho)(R\varphi_\varrho, \varphi_\nu) = \sum_{\varrho=1}^{\infty} (S\varphi_\mu, \varphi_\varrho)(\varphi_\varrho, R\varphi_\nu) = (S\varphi_\mu, R\varphi_\nu),$$

also lautet die Vert.-Bedingung einfach so:

$$(R\varphi_\mu, S\varphi_\nu) = (S\varphi_\mu, R\varphi_\nu).$$

Wenn nun alle $R(S\varphi_\mu)$, $S(R\varphi_\mu)$ sinnvoll sind, so können wir hierfür schreiben:

$$(SR\varphi_\mu, \varphi_\nu) = (RS\varphi_\mu, \varphi_\nu), \qquad ((RS - SR)\varphi_\mu, \varphi_\nu) = 0,$$

und da die φ_ν ein vollst. norm. orth. System bilden, $(RS - SR)\varphi_\mu = 0$. Wenn also irgendein abg. lin. O. T Forts. von $i(RS - SR)$ ist, so gilt $Tf = 0$ für alle $f = \varphi_\mu$, also auch für alle Linearaggregate endlich vieler — d. h. in einer überall dichten Menge. Da nun T abg. ist, ist danach Tf stets sinnvoll, und zwar $= 0$, also $T = 0$. Andererseits ist $i(RS - SR)$ offenbar H. (es ist für alle φ_μ sinnvoll, d. h. sein Definitionsbereich spannt die abg. lin. M. $\mathfrak{H}$ auf), somit können wir $T = \widetilde{i(RS - SR)}$ bilden, und dieses muß $= 0$ sein.

In diesem Falle können wir also mit einigem Recht von einer Vert. von R, S reden — insbesondere für R, S aus $\mathfrak{B}$ (beschr.!) trifft dies zu, da aber $i(RS - SR)$ dann sogar überall sinnvoll, also selbst abg. lin. ist, ist $i(RS - SR) = 0$, $RS = SR$. Wie steht es aber bei beliebigen R, S, für die die $R(S\varphi_\mu)$, $S(R\varphi_\mu)$ nicht alle sinnvoll sein müssen? Inwieweit bedeutet dann $(R\varphi_\mu, S\varphi_\nu) = (S\varphi_\mu, R\varphi_\nu)$ ($\mu, \nu = 1, 2, \ldots$, im vollst. norm. orth. System $\varphi_1, \varphi_2, \ldots$ mögen R, S Matrizen haben; vgl. Anm. [78])) eine vernünftige Vert. von R, S?

Im folgenden soll an einem Gegenbeispiel gezeigt werden, daß die genannte Relation in dieser Allgemeinheit kaum von Bedeutung sein kann.

Zunächst sei aber noch erwähnt, auf welchen Begriff der Normalität (für konj. Paare R, R^*) dieser Vertauschbarkeitsbegriff (für H. O.) führt. Wir müssen dann $S_1 = \frac{R + R^*}{2}$, $S_2 = \frac{R - R^*}{2i}$ bilden, und

$$(S_1\varphi_\mu, S_2\varphi_\nu) = (S_2\varphi_\mu, S_1\varphi_\nu)$$

verlangen, oder, wie man leicht umrechnet,

$$(R\varphi_\mu, R\varphi_\nu) = (R^*\varphi_\mu, R^*\varphi_\nu).$$

Dies bedeutet offenbar:

$$|Rf| = |R^* f|$$

für alle Linearaggregate endlich vieler φ_μ. Daß das für die Normalität notwendig ist, wissen wir aus Satz 21.

2. Sei $\varphi_1, \varphi_2, \ldots$ ein vollst. norm. orth. System. Wir definieren zwei Matrizen $\{a_{\mu\nu}\}, \{b_{\mu\nu}\}$ so: wenn nicht μ und ν gleich $2n-1$ oder $2n$ (für dasselbe $n = 1, 2, \ldots$) ist, so sei $a_{\mu\nu} = b_{\mu\nu} = 0$; ferner

$$a_{2n-1,2n-1} = a_{2n,2n-1} = a_{2n-1,2n} = a_{2n,2n} = n;$$

$$b_{2n-1,2n-1} = b_{2n,2n} = n, \quad b_{2n-1,2n} = -b_{2n,2n-1} = in.$$

Die zu den (offenbar H. quadrierbaren) Matrizen $\{a_{\mu\nu}\}$ bzw. $\{b_{\mu\nu}\}$ (und $\{\varphi_1, \varphi_2, \ldots\}$) gehörigen abg. lin. H. O. seien R bzw. S (vgl. Anm. [73])).

Durch

$$\psi_{2n-1} = \frac{1}{\sqrt{2}}\varphi_{2n-1} + \frac{1}{\sqrt{2}}\varphi_{2n}, \quad \psi_{2n} = \frac{1}{\sqrt{2}}\varphi_{2n-1} - \frac{1}{\sqrt{2}}\varphi_{2n};$$

$$\chi_{2n-1} = \frac{1}{\sqrt{2}}\varphi_{2n-1} + \frac{i}{\sqrt{2}}\varphi_{2n}, \quad \chi_{2n} = \frac{1}{\sqrt{2}}\varphi_{2n} \quad - \frac{i}{\sqrt{2}}\varphi_{2n}$$

sind ebenfalls vollst. norm. Orth. Systeme $\psi_1, \psi_2, \ldots$ und $\chi_1, \chi_2, \ldots$ definiert, und zwar gilt

$$R\psi_{2n-1} = 2n\cdot\psi_{2n-1}, \quad R\psi_{2n} = 0; \quad S\chi_{2n-1} = 2n\cdot\chi_{2n-1}, \quad S\chi_{2n} = 0,$$

dabei sind die R, S abg. lin. In diesen vollst. norm. orth. Systemen haben also R bzw. S Diagonalmatrizen, also sind beide hypermax.[74]). Insbesondere hat Rf für $f = \sum_{\nu=1}^{\infty} y_\nu \psi_\nu$ ($\sum_{\nu=1}^{\infty} |y_\nu|^2$ endlich) dann und nur dann Sinn, wenn $\sum_{n=1}^{\infty} 4n^2 \cdot |y_{2n-1}|^2$ endlich ist; für $f = \sum_{\nu=1}^{\infty} x_\nu \varphi_\nu$ ($\sum_{\nu=1}^{\infty} |x_\nu|^2$ endlich) ist aber $y_{2n-1} = \frac{1}{\sqrt{2}} x_{2n-1} + \frac{1}{\sqrt{2}} x_{2n}$, daher ist die Endlichkeit von $\sum_{n=1}^{\infty} n^2 |x_{2n-1} + x_{2n}|^2$ charakteristisch. Ebenso zeigt man, daß die Sinnvollheit von Sf für $f = \sum_{\nu=1}^{\infty} z_\nu \chi_\nu$ ($\sum_{\nu=1}^{\infty} |z_\nu|^2$ endlich) durch die Endlichkeit von $\sum_{n=1}^{\infty} 4n^2 |z_{2n-1}|^2$, also für $f = \sum_{\nu=1}^{\infty} x_\nu \varphi_\nu$, wo $z_{2n-1} = \frac{1}{\sqrt{2}} x_{2n-1} - \frac{i}{\sqrt{2}} x_{2n}$ ist, durch diejenige von $\sum_{n=1}^{\infty} n^2 |x_{2n-1} - i x_{2n}|^2$ gekennzeichnet ist. Die Endlichkeit beider Summen ist, wie man leicht einsieht, mit der von

[74]) Allgemein folgt aus $T\xi_n = \alpha_n \xi_n$ ($n = 1, 2, \ldots$, T ein abg. lin. O., $\xi_1, \xi_2, \ldots$ ein vollst. norm. orth. System) die Hypermax. von T. In der Tat umfassen die abg. lin. M. $\mathfrak{E}, \mathfrak{F}$ (vgl. E. Kap. V) alle $(T \pm i1)\xi_n = (\alpha_n \pm i)\xi_n$, also alle ξ_n. Daher sind beide gleich $\mathfrak{H}$.

$\sum_{n=1}^{\infty} n^2(|x_{2n-1}|^2 + |x_{2n}|^2)$ gleichwertig. Wenn nun noch $R(Sf)$ und $S(Rf)$ sinnvoll sein sollen, so ist $Rf = \sum_{\nu=1}^{\infty} u_\nu \varphi_\nu$, $Sf = \sum_{\nu=1}^{\infty} v_\nu \varphi_\nu$ mit

$$u_{2n-1} = n \cdot x_{2n-1} + n \cdot x_{2n}, \quad u_{2n} = n \cdot x_{2n-1} + n \cdot x_{2n};$$
$$v_{2n-1} = n \cdot x_{2n-1} - in \cdot x_{2n}, \quad v_{2n} = in \cdot x_{2n-1} + n \cdot x_{2n}$$

zu betrachten: es muß $\sum_{n=1}^{\infty} n^2 |v_{2n-1} + v_{2n}|^2$ und $\sum_{n=1}^{\infty} n^2 |u_{2n-1} - i u_{2n}|^2$ endlich sein, d. h. $\sum_{n=1}^{\infty} n^4 |x_{2n-1} + x_{2n}|^2$ und $\sum_{n=1}^{\infty} n^4 |x_{2n-1} - i x_{2n}|^2$. Man überlegt sich leicht, daß die Endlichkeit dieser Summen derjenigen von $\sum_{n=1}^{\infty} n^4 (|x_{2n-1}|^2 + |x_{2n}|^2)$ gleichwertig ist.

Nun bestimmen wir für diese f $i(RS - SR)f = \sum_{\nu=1}^{\infty} w_\nu \varphi_\nu$, es ergibt sich:

$$z_{2n-1} = -2n^2 \cdot x_{2n-1}, \quad z_{2n} = 2n^2 \cdot x_{2n}.$$

Daher ist dieses $i(RS - SR)$ ein abg. lin. H. O., und hat eine Diagonalmatrix im System $\varphi_1, \varphi_2, \dots$ — also ist es hypermax. (vgl. Anm. [74])). Daß es $\neq 0$ ist, ist klar: wie man sieht, folgt sogar aus $f \neq 0$ $i(RS - SR)f \neq 0$.

Von einer Vert. von R, S kann also keine Rede sein. Trotzdem werden wir ein vollst. norm. orth. System $\omega_1, \omega_2, \dots$ auffinden, in dem R, S beide Matrizen besitzen, und die Vert.-Relationen des § 1 erfüllt sind.

3. Daß R und S in einem vollst. norm. orth. System $\omega_1, \omega_2, \dots$ Matrizen $\{a_{\mu\nu}\}$ bzw. $\{b_{\mu\nu}\}$ besitzen, besagt folgendes (vgl. Anm. [73])): es muß jedenfalls $a_{\mu\nu} = (R\omega_\mu, \omega_\nu)$, $b_{\mu\nu} = (S\omega_\mu, \omega_\nu)$ sein, die elementaren O. dieser Matrizen, R' bzw. S', sind die auf die Menge der $\omega_1, \omega_2, \dots$ eingeschränkten O. R bzw. S, die abg. lin. O. sind $\tilde{R}'$ bzw. $\tilde{S}'$ — und es wird verlangt $\tilde{R}' = R$, $\tilde{S}' = S$. D. h.: zu jedem f mit sinnvollem Rf bzw. Sf soll es eine Folge $f_1, f_2, \dots$ aus der lin. M. (nicht der abg. lin. M.!) der $\omega_1, \omega_2, \dots$ geben, mit $f_n \to f$ und $Rf_n \to Rf$ bzw. $Sf_n \to Sf$ (wenn Rf, Sf beide Sinn haben, können es zwei verschiedene Folgen sein!). Wenn $\omega_1, \omega_2, \dots$ durch das E. Schmidtsche Orthogonalisationsverfahren [75]) aus einer überall dichten Folge $g_1, g_2, \dots$ entsteht, so ist dies gewiß der Fall, wenn die soeben genannten $f_1, f_2, \dots$ mit $f_n \to f$ und $Rf_n \to Rf$ bzw. $Sf_n \to Sf$ als Teilfolge von $g_1, g_2, \dots$ gewählt werden können. Und wenn stets $(Rg_p, Sg_q) = (Sg_p, Rg_q)$ gilt, so gilt dies auch für die $\omega_1, \omega_2, \dots$ (die Linearaggregate der g sind): $(R\omega_\mu, S\omega_\nu) = (S\omega_\mu, R\omega_\nu)$, d. h. R, S erfüllen die Vert.-Relation des § 1. Es sei noch betont, daß wir die Sinnvollheit der Rg_p, Sg_p voraussetzten, aber nicht die der $R(Sg_p), S(Rg_p)$.

[75]) Vgl. E. Schmidt, am in Anm. [56]) a. O. Siehe auch E. § 1, Satz 8.

Sei $\bar{g}_1, \bar{g}_2, \ldots$ die Menge aller Linearaggregate endlich vieler $\varphi_1, \varphi_2, \ldots$ mit rationalen Koeffizienten, als Folge geschrieben. Da R, S im System $\varphi_1, \varphi_2, \ldots$ Matrizen haben, sieht man sofort, daß die früher erwähnten Folgen $f_1, f_2, \ldots$ aus $\bar{g}_1, \bar{g}_2, \ldots$ ausgewählt werden können (die Rationalitätsbedingung für die Koeffizienten stört nicht). Seien nun $h_1, h_2, \ldots$ so gewählt, daß alle Rh_n, Sh_n Sinn haben, und (wie immer in $\mathfrak{H}$, stark)

$$h_n \to 0, \quad Rh_{2n-1} \to 0, \quad Sh_{2n} \to 0 \quad (\text{für } n \to \infty)$$

gilt. Dann sieht man, daß die obigen Folgen $f_1, f_2, \ldots$ stets auch als Teilfolgen von $\bar{g}_1 + h_1, \bar{g}_2 + h_2, \ldots, \bar{g}_1 + h_3, \bar{g}_2 + h_4, \ldots$ gewählt werden können. Wir schreiben für $\bar{g}_1, \bar{g}_1, \bar{g}_2, \bar{g}_2, \ldots$ kürzer $\tilde{g}_1, \tilde{g}_2, \ldots$, und wählen $g_1, g_2, \ldots$ als $\tilde{g}_1 + h_1, \tilde{g}_2 + h_2, \ldots$. Daher ist alles erreicht, wenn noch $(Rg_p, Sg_q) = (Sg_p, Rg_q)$ gilt — dies gilt es durch geeignete Wahl der $h_1, h_2, \ldots$ zu erzwingen (die Wahl der $h_1, h_2, \ldots$ ist ja noch bis zu einem gewissen Grade frei, während die $\tilde{g}_1, \tilde{g}_2, \ldots$ bereits festgelegt sind).

Wir indizieren die Folge $\varphi_1, \varphi_2, \ldots$ folgendermaßen um: Wir teilen sie in Paare $\varphi_{2n-1}, \varphi_{2n}$ ein und ersetzen die einfache Indizierung $n = 1, 2, \ldots$ durch eine dreifache mit p, q, μ (alle $= 1, 2, \ldots$), und zwar sollen $\varphi_{2n-1}, \varphi_{2n}$ bzw. $\varphi'_{pq\mu}; \varphi''_{pq\mu}$ heißen. Das zu p, q, μ gehörige n heiße $n(pq\mu)$, wir wollen es so einrichten, daß stets $n(pq\mu) \geqq \mu^2$ ist. Nun wird h_p definiert:

$$h_p = \sum_{q,\mu=1}^{\infty} \frac{1}{pq\cdot(n(pq\mu))^{\frac{1}{2}}} \cdot \varphi'_{pq\mu} + \sum_{r=1}^{p-1} x_{rp} \cdot v_{rp},$$

dabei ist

$$v_{rp} = \left\{ \begin{array}{ll} \varphi'_{r,p,\varrho(rp)} - \varphi''_{r,p,\varrho(rp)}, & \text{für ungerades } p \\ \varphi'_{r,q,\varrho(rp)} - i\,\varphi''_{r,p,\varrho(rp)}, & \text{für gerades } p \end{array} \right| .$$

Über die x_{rp} und $\varrho(rp)$ $(r < p)$ verfügen wir noch später. Betrachten wir zuerst die Summen $\xi_p = \sum_{q,\mu=1}^{\infty} \frac{1}{pq\cdot(n(pq\mu))^{\frac{1}{2}}} \varphi'_{pq\mu}$. Sie sind konvergent, ihre Absolutwertquadratsummen sind

$$\sum_{q,\mu=1}^{\infty} \frac{1}{p^2 q^2 (n(pq\mu))^{\frac{1}{2}}} \leqq \sum_{q,\mu=1}^{\infty} \frac{1}{p^2 q^2 \mu^{\frac{1}{2}}} = \frac{1}{p^2}\left(\sum_{q=1}^{\infty}\frac{1}{q^2}\right)\left(\sum_{\mu=1}^{\infty}\frac{1}{\mu^{\frac{1}{2}}}\right),$$

daher streben sie mit $p \to \infty$ gegen 0. Da auch die Summen

$$\sum_{q,\mu=1}^{\infty} \frac{1}{p^2 q^2 (n(pq\mu))^{\frac{1}{2}}} (n(pq\mu))^2 = \sum_{q,\mu=1}^{\infty} \frac{1}{p^2 q^2 n(pq\mu)} \leqq \sum_{q,\mu=1}^{\infty} \frac{1}{p^2 q^2 \mu^2}$$

$$= \frac{1}{p^2}\left(\sum_{q=1}^{\infty}\frac{1}{q^2}\right)\left(\sum_{\mu=1}^{\infty}\frac{1}{\mu^2}\right)$$

endlich sind (sie streben sogar mit $p \to \infty$ gegen 0), sind auch die $R\xi_p$, $S\xi_p$ sinnvoll. $\sum_{r=1}^{p-1} x_{rp} \cdot v_{rp}$ ist eine endliche Summe, daher ist darauf R und S anwendbar; sie strebt mit $p \to \infty$ gegen 0, wenn $\sum_{r=1}^{p-1} |x_{rp}|^2 \to 0$, was z. B. für $|x_{rp}| \leqq \frac{1}{rp}$ gewiß der Fall ist. Ferner ist für ungerades bzw. gerades p $Rv_{rp} = 0$ bzw. $Sv_{rp} = 0$, d. h. R bzw. S auf $\sum_{r=1}^{p-1} x_{rp} v_{rp}$ angewandt ergibt 0. Zusammenfassend sehen wir: Rh_p, Sh_p sind stets sinnvoll, $h_p \to \infty$, $Rh_p \to 0$ für ungerade p, $Sh_p \to 0$ für gerade p.

Es gilt nun noch $(Rg_p, Sg_q) - (Sg_p, Rg_q) = 0$ durch geeignete Wahl der x_{rp}, $\varrho(rp)$ zu sichern (man sieht, daß es genügt, nur $p < q$ zu betrachten, was geschehen soll) — natürlich unter Wahrung von $|x_{rp}| \leqq \frac{1}{rp}$. Wir setzen in die obige Formel

$$g_p = \tilde{g}_p + \sum_{t,\mu=1}^{\infty} \frac{1}{pt(n(pt\mu))^{\frac{3}{2}}} \cdot \varphi'_{pt\mu} + \sum_{r=1}^{p-1} x_{rp} \cdot v_{rp},$$

$$g_q = \tilde{g}_q + \sum_{\mu,\nu=1}^{\infty} \frac{1}{qu(n(qu\nu))^{\frac{3}{2}}} \cdot \varphi'_{qu\nu} + \sum_{s=1}^{q-1} x_{sq} \cdot v_{sq}$$

ein, und rechnen den Ausdruck aus. Es wird, unter Berücksichtigung der bestehenden Orthogonalitäten (es ist $p < q$!)

$$\begin{aligned}(\{SR - RS\}\tilde{g}_p, \tilde{g}_q) + (\{SR - RS\}\tilde{g}_p, \xi_q) + (\xi_p, \{RS - SR\}\tilde{g}_q) \\ + \sum_{s=1}^{q-1} \bar{x}_{sq}(\{SR - RS\}\tilde{g}_p, v_{sq}) + \sum_{r=1}^{p-1} x_{rp}(v_{rp}, \{RS - SR\}\tilde{g}_q) \\ + \frac{\bar{x}_{pq}}{pq(n(p,q,\varrho(pq)))^{\frac{3}{2}}}(\{SR - RS\}\varphi'_{p,q,\varrho(pq)}, v_{pq}) = 0.\end{aligned}$$

Wir setzen zur Abkürzung $i(RS - SR) = T$. Ferner führen wir die folgende Bezeichnung ein: $\tilde{g}_l$ ist ein Linearaggregat endlich vieler φ_n, also endlich vieler $\varphi'_{pq\mu}$ und $\varphi''_{pq\mu}$, das größte dabei vorkommende μ sei $\mu(l)$. Wir schreiben nun vor, daß stets $\varrho(rt) > \text{Max}(\mu(1), \ldots, \mu(t-1))$ sein soll — dann fällt in unserem Ausdruck das vierte Glied identisch fort. Es wird:

$$\begin{aligned}(T\tilde{g}_p, \tilde{g}_q) + (T\tilde{g}_p, \xi_q) + (\xi_p, T\tilde{g}_q) + \sum_{r=1}^{p-1} x_{rp}(v_{rp}, T\tilde{g}_q) \\ + \frac{\bar{x}_{pq}}{pq(n(p,q,\varrho(pq)))^{\frac{3}{2}}}(T\varphi'_{p,q,\varrho(pq)}, v_{pq}) = 0.\end{aligned}$$

Wenn wir noch $(T\varphi'_{p,q,\varrho(pq)}, v_{pq}) = -2(n(p,q,\varrho(pq)))^2$ beachten, so wird

$$\begin{aligned}\bar{x}_{pq} = -\frac{1}{2(n(p,q,\varrho(pq)))^{\frac{1}{2}}}\Big[(T\tilde{g}_p, \tilde{g}_q) + (T\tilde{g}_p, \xi_q) + (\xi_p, T\tilde{g}_q) \\ + \sum_{r=1}^{p-1} x_{rp}(v_{rp}, T\tilde{g}_q)\Big].\end{aligned}$$

Hier ist x_{pq} mit Hilfe der x_{rp} und v_{rp}, also der $\varrho(rp)$, mit $r = 1, \ldots, p-1$ sowie von $\varrho(pq)$ ausgedrückt. Dabei ist $\varrho(pq)$ willkürlich. Wenn wir also die x_{rs}, $\varrho(rs)$ nach wachsendem $r+s$ anordnen, so liegt hier eine Rekursionsgleichung für die x_{rs} vor — und die $\varrho(rs)$ sind dabei beliebig. Es gilt nur noch $|x_{pq}| \leqq \frac{1}{pq}$ und $\varrho(pq) \geqq \operatorname{Max}(\mu(1), \ldots, \mu(q-1))$ einzuhalten. Wir wählen die $\varrho(rs)$ auch in der genannten Reihenfolge, dann haben wir (C ist der Absolutwert des Inhalts von $[\ldots]$ in der Formel für $\bar{x}_{pq}$, hängt also nur von x_{rs}, $\varrho(rs)$ mit $r+s < p+q$ ab, von $\varrho(pq)$ aber nicht!)

$$|x_{pq}| = \frac{C}{2\left(n(p,q,\varrho(pq))\right)^{\frac{1}{2}}} \leqq \frac{C}{2\varrho(pq)}.$$

Somit genügt es, $\varrho(pq) \geqq \frac{1}{2} Cpq$ und $\geqq \operatorname{Max}(\mu(1), \ldots, \mu(q-1))$ zu wählen, und wir sind am Ziele.

Wir haben jetzt die Folgen $h_1, h_2, \ldots$ sowie $g_1, g_2, \ldots$ und das System $\omega_1, \omega_2, \ldots$, auf das es uns ankam: damit ist die gewünschte Konstruktion durchgeführt.

(Eingegangen am 20. 2. 1929.)

Bemerkung über die konformen Abbildungen konvexer Gebiete.

Von

Tibor Radó[1]) in Szeged, z. Z. München.

Die Funktion $w = f(z)$ bilde das Innere des Einheitskreises $|z| < 1$ konform auf ein *konvexes* Gebiet G der w-Ebene ab. Nach Study[2]) werden dann auch die konzentrischen Kreisflächen $|z| < r < 1$ auf konvexe Gebiete abgebildet. Ich möchte in dieser Note einen sehr elementaren Beweis dieses schönen Satzes mitteilen, wobei derselbe als unmittelbare Folgerung des Schwarzschen Lemmas erscheint[3]).

Beim Beweise dürfen wir offenbar

$$f(0) = 0 \tag{1}$$

voraussetzen. Wir bezeichnen die Kreisscheibe $|z| < r < 1$ mit K_r, ihr Bildgebiet mit G_r, und wir wählen in G_r irgend zwei verschiedene Punkte w_1, w_2. Zu zeigen ist, daß jeder auf der Verbindungsstrecke von w_1 und w_2 gelegene Punkt w_0 ebenfalls in G_r enthalten ist; falls wir die inverse Funktion von $f(z)$ mit $F(w)$ bezeichnen, so ist diese Behauptung der Ungleichung

$$|F(w_0)| < r \tag{2}$$

gleichwertig, die ja eben die Tatsache ausdrückt, daß der Urbildpunkt von w_0 in K_r liegt[4]).

---

[1]) International Research Fellow.

[2]) Study, Konforme Abbildung einfach-zusammenhängender Bereiche, S. 110 ff., B. G. Teubner, Leipzig und Berlin, 1913.

[3]) Für den üblichen Beweis vgl. etwa die Darstellung bei Carathéodory, Sur la représentation conforme des polygones convexes, Annales de la Société scientifique de Bruxelles 37 (1913), S. 1—10.

[4]) Um vom Urbildpunkte von w_0 sprechen zu können, muß w_0 in G liegen; dies ist aber gewiß der Fall, denn G ist konvex, und w_0 liegt auf der Verbindungsstrecke von zwei Punkten von G.

Der Punkt w_0 kann, weil er auf der Verbindungsstrecke von w_1 und w_2 liegt, in der Form

$$(3) \qquad w_0 = t w_1 + (1-t) w_2, \qquad t \text{ reell und } 0 < t < 1$$

dargestellt werden. Es seien z_1 und z_2 die Urbildpunkte von w_1 und w_2, wobei die Bezeichnung so gewählt werde, daß $|z_1| \leqq |z_2|$ ausfällt. Dann sind z_1 und z_2 zwei verschiedene Punkte in K_r, und es ist gewiß $z_2 \neq 0$, weil sonst $z_1 = 0 = z_2$ wäre. Wir haben also:

$$(4) \qquad |z_1| \leqq |z_2| < r, \qquad z_2 \neq 0.$$

Wir bilden nun die, wegen (4) für $|z| < 1$ reguläre, Funktion

$$(5) \qquad \varphi(z) = t f\left(\frac{z_1}{z_2} z\right) + (1-t) f(z).$$

In der w-Ebene dargestellt, fallen die Punkte $f\left(\frac{z_1}{z_2} z\right)$ und $f(z)$ in G, $\varphi(z)$ ist aber, wegen $0 < t < 1$, ein Punkt auf der Verbindungsstrecke dieser beiden Punkte, also ebenfalls ein Punkt in G, da G konvex ist. Wir dürfen also $\varphi(z)$ in die inverse Funktion $F(w)$ einsetzen und erhalten eine für $|z| < 1$ reguläre Funktion

$$(6) \qquad \psi(z) = F(\varphi(z)),$$

die wegen (1) und (5) für $z = 0$ verschwindet und für $|z| < 1$ dem absoluten Betrage nach kleiner als Eins ist, weil nämlich F nur solche Werte annimmt, die absolut < 1 sind. Es erfüllt also $\psi(z)$ die Voraussetzungen des Schwarzschen Lemmas und wir haben daher

$$|\psi(z)| \leqq |z| \quad \text{für} \quad |z| < 1.$$

Setzen wir hier $z = z_2$, so ergibt sich, da $|z_2| < r$ und, wegen (6), (5) und (3), $\psi(z_2) = F(w_0)$ ist, gerade die zu beweisende Ungleichung (2) [5].

[5]) Herrn Carathéodory habe ich für freundliche Ratschläge zur endgültigen Darstellung dieses Beweises zu danken.

(Eingegangen am 24. 7. 1929.)

Über die Existenz von Repräsentantenbereichen in der Theorie der Abbildung durch Paare von Funktionen zweier komplexen Veränderlichen.

Von

Stefan Bergmann in Berlin.

In Analogie zu dem Abbildungssatz der Theorie der Funktionen einer komplexen Veränderlichen hat Poincaré das folgende Problem[1]) gestellt: Es sind zwei einfach zusammenhängende Bereiche $\mathfrak{B}$ und $\mathfrak{B}^*$ des P_4[2]) gegeben; es ist zu entscheiden, ob es ein Funktionenpaar

$$(1) \qquad X^*(X, Z), \quad Z^*(X, Z)$$

von zwei komplexen Veränderlichen gibt, das $\mathfrak{B}$ auf $\mathfrak{B}^*$ abbildet[3]). Poincaré zeigte, daß eine solche Abbildung nicht stets möglich ist. Reinhardt[4]) hat nachher mit ganz anderen Mitteln bewiesen, daß im allgemeinen eine Abbildung von zwei Kreisbereichen aufeinander nicht möglich ist[5]).

Unter der Voraussetzung, daß die Abbildung von $\mathfrak{B}$ auf $\mathfrak{B}^*$ möglich ist, wurde in der Arbeit: „Über unendliche Hermitesche Formen, die zu einem Bereiche gehören, nebst Anwendungen auf Fragen der Abbildung durch Funktionen von zwei komplexen Veränderlichen"[6]) folgendes gezeigt: Man kann die Koeffizienten der Funktionenelemente des Abbildungspaares

[1]) H. Poincaré, Les fonctions analytiques de deux variables et la représentation conforme, Rendiconti di Circolo Matematico di Palermo **23** (1907).

[2]) Mit P_n werden wir den n-dimensionalen Raum bezeichnen.

[3]) Darunter versteht man folgendes: Während X, Z alle inneren Punkte von $\mathfrak{B}$ durchläuft, durchläuft X^*, Z^* alle inneren Punkte von $\mathfrak{B}^*$.

[4]) K. Reinhardt, Über Abbildungen durch analytische Funktionen zweier Veränderlichen, Math. Annalen **83** (1921), S. 211–255.

[5]) Weitere wertvolle Beiträge zur Theorie stammen von Behnke, Blaschke, Carathéodory und Kritikos.

[6]) Math. Zeitschr. **29** (1929), S. 641–677. Wir werden diese Arbeit im folgenden als Arbeit H zitieren.

in einfacher Weise durch die Werte (im Koordinatenanfangspunkte) der Orthogonalfunktionen $\Omega_s(X, Z)$ und $\varphi_s(X^*, Z^*)$ und deren Ableitungen ausdrücken [7]). $\Omega_s(X, Z)$ und $\varphi_s(X^*, Z^*)$ sind die zu $\mathfrak{B}$ bzw. $\mathfrak{B}^*$ gehörigen vollständigen Systeme von Orthogonalfunktionen [8]). Dadurch erhält man — zumindest im Prinzip — eine Methode, um zu entscheiden, ob eine Abbildung von $\mathfrak{B}$ auf $\mathfrak{B}^*$ möglich ist oder nicht.

Ein anderer Weg, ein solches Kriterium zu gewinnen, ist der folgende: Faßt man die Gesamtheit derjenigen Bereiche, die durch (in bezug auf den Punkt $0, 0$) normierte Abbildungen [9]) auseinander hervorgehen, zu einer Klasse zusammen, so entsteht das Problem:

1. in einer Klasse einen ausgezeichneten Repräsentantenbereich $\mathfrak{A}$ auszusondern [10]), und

2. ein Verfahren anzugeben, welches das Abbildungspaar von $\mathfrak{B}$ auf $\mathfrak{A}$ zu berechnen erlaubt.

Behandelt man zunächst das zweite Problem, so ist diese Frage mit der Aufgabe identisch, in der Klasse ein normiertes Kovariantenpaar [11]) $v(X, Z)$, $w(X, Z)$ zu bestimmen [12]).

---

[7]) Man macht dabei die Voraussetzung, daß der Koordinatenanfangspunkt ein innerer Punkt von $\mathfrak{B}$ und $\mathfrak{B}^*$ ist.

[8]) Unter einem zu $\mathfrak{B}$ gehörigen Orthogonalfunktionensystem versteht man ein Funktionensystem, das die Orthogonalrelation $\iiint\!\!\int_{\mathfrak{B}} \Omega_s(X, Z)\, \overline{\Omega_k(X, Z)}\, d\omega = \begin{matrix} 1 \\ 0 \end{matrix} \begin{pmatrix} s = k \\ s \neq k \end{pmatrix}$ befriedigt. Unter $\iiint\!\!\int_{\mathfrak{B}}$ wird stets ein uneigentliches Integral $\lim_{n \to \infty} \iiint\!\!\int_{\mathfrak{B}_n}$ verstanden, wobei $\mathfrak{B}_n$ eine Folge ganz im Innern von $\mathfrak{B}$ gelegener Bereiche, die gegen $\mathfrak{B}$ konvergieren, ist.

[9]) Ein Funktionenpaar V, W heißt in bezug auf den Punkt $0, 0$ normiert, falls

$$V(0,0) = 0, \qquad \left[\frac{\partial V(X,Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = 1, \qquad \left[\frac{\partial V(X,Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = 0,$$

$$W(0,0) = 0, \qquad \left[\frac{\partial W(X,Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = 0, \qquad \left[\frac{\partial W(X,Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = 1$$

ist.

[10]) Es darf selbstverständlich in einer Klasse nur ein Bereich $\mathfrak{A}$ existieren.

[11]) Darunter wird folgendes verstanden: Bedeuten $\mathfrak{B}$ und $\mathfrak{B}^{**}$ zwei Bereiche derselben Klasse; X, Z bzw. X^{**}, Z^{**} die sich entsprechenden (d. h. bei der Abbildung ineinander übergehenden) Punkte von $\mathfrak{B}$ und $\mathfrak{B}^{**}$; $v(X, Z)$, $w(X, Z)$ das zu $\mathfrak{B}$, $v^{**}(X^{**}, Z^{**})$, $w^{**}(X^{**}, Z^{**})$ das zu $\mathfrak{B}^{**}$ gehörige Paar, so soll

$$v(X, Z) = v^{**}(X^{**}, Z^{**}), \qquad w(X, Z) = w^{**}(X^{**}, Z^{**})$$

sein.

[12]) Bei $v(X, Z)$ und $w(X, Z)$ sind in $\mathfrak{B}$ nur außerordentliche singuläre Stellen zugelassen. Die Funktionaldeterminante $\dfrac{\partial(v, w)}{\partial(X, Z)}$ darf nicht identisch verschwinden.

Der Zweck der vorliegenden Arbeit besteht darin, zu zeigen, daß man mit Hilfe des vollständigen Orthogonalfunktionensystems $\Omega_s(X, Z)$, das zu $\mathfrak{B}$ gehört, ein derartiges Kovariantenpaar bestimmen kann; wie im § 1 gezeigt wird, *bildet*

$$
(2)\qquad
\begin{aligned}
v(X, Z) &= -c\,\frac{\begin{vmatrix} \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}_s & \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}'_{sX} & \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}'_{sZ} \\ \Sigma\,\Omega_s\,\overline{\Omega}_s & \Sigma\,\Omega_s\,\overline{\Omega}'_{sX} & \Sigma\,\Omega_s\,\overline{\Omega}'_{sZ} \\ \Sigma\,\Omega'_{sZ}\,\overline{\Omega}_s & \Sigma\,\Omega'_{sZ}\,\overline{\Omega}'_{sX} & \Sigma\,\Omega'_{sZ}\,\overline{\Omega}'_{sZ} \end{vmatrix}}{\Sigma\,\Omega_s(X,Z)\,\overline{\Omega}_s} \\
w(X, Z) &= c\,\frac{\begin{vmatrix} \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}_s & \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}'_{sX} & \Sigma\,\Omega_s(X,Z)\,\overline{\Omega}'_{sZ} \\ \Sigma\,\Omega_s\,\overline{\Omega}_s & \Sigma\,\Omega_s\,\overline{\Omega}'_{sX} & \Sigma\,\Omega_s\,\overline{\Omega}'_{sZ} \\ \Sigma\,\Omega'_{sX}\,\overline{\Omega}_s & \Sigma\,\Omega'_{sX}\,\overline{\Omega}'_{sX} & \Sigma\,\Omega'_{sX}\,\overline{\Omega}'_{sZ} \end{vmatrix}}{\Sigma\,\Omega_s(X,Z)\,\overline{\Omega}_s}
\end{aligned}
$$

ein derartiges Kovariantenpaar. Es bedeuten darin:

$$
(3)\qquad \Omega_s = \Omega_s(0,0),\qquad \Omega'_{sX} = \left[\frac{\partial\,\Omega_s(X,Z)}{\partial X}\right]_{\substack{X=0\\ Z=0}},\qquad \Omega'_{sZ} = \left[\frac{\partial\,\Omega_s(X,Z)}{\partial Z}\right]_{\substack{X=0\\ Z=0}},
$$

c eine Konstante und

$$
(4)\qquad \Sigma \equiv \sum_{s=1}^{s=\infty}.
$$

Um die Möglichkeit einer Abbildung von $\mathfrak{B}$ auf $\mathfrak{B}^*$ festzustellen, bildet man das zu $\mathfrak{B}$ bzw. $\mathfrak{B}^*$ gehörige Kovariantenpaar v, w bzw. v^*, w^*. Durch ein solches wird $\mathfrak{B}$ auf den Repräsentantenbereich seiner Klasse, den Bereich $\mathfrak{A}$, $\mathfrak{B}^*$ auf $\mathfrak{A}^*$ abgebildet. *Eine Abbildung von $\mathfrak{B}$ auf $\mathfrak{B}^*$ ist dann und nur dann möglich, wenn $\mathfrak{A}$ mit $\mathfrak{A}^*$ übereinstimmt.*

Im § 2 wird auf die Frage der Bildung eines vollständigen Orthogonalfunktionensystems zu einem Bereiche eingegangen. In der Arbeit: „Zwei Sätze über Funktionen von zwei komplexen Veränderlichen" wurde mit Hilfe des Auswahlprinzips die Existenz eines vollständigen Orthogonalfunktionensystems zu jedem Bereich $\mathfrak{B}$ bewiesen. Da es aber wünschenswert erscheint, ein konstruktives Verfahren zur Herstellung von vollständigen Systemen anzugeben, wird im § 2 für eine Reihe von einfach zusammenhängenden Bereichen (die mit $\mathfrak{R}$ bezeichnet werden) der Approximationssatz bewiesen: *Man kann jede im abgeschlossenen Bereiche $\mathfrak{R}$ reguläre Funktion gleichmäßig in jedem ganz im Innern von $\mathfrak{R}$ gelegenen Teilbereiche $\mathfrak{R}'$ durch Polynome approximieren*[13]).

Eine ausführliche Definition der Bereiche $\mathfrak{R}$ befindet sich im Text auf S. 443. Aus dem Approximationssatz folgt (genau wie im Falle einer

[13]) Es ist eine Verallgemeinerung des bekannten Rungeschen Satzes: C. Runge, Zur Theorie der eindeutigen analytischen Funktionen, Acta Mathematica 6 (1884), S. 229.

komplexen Veränderlichen), daß das durch Orthogonalisierung von

$$(5) \qquad 1, X, Z, X^2, XZ, Z^2, \ldots$$

erzeugte System in $\mathfrak{R}$ vollständig ist.

Bei dieser Gelegenheit benutzen wir eine Integraldarstellung der Funktion in $\mathfrak{R}$. Diese Darstellung erlaubt zu ersehen, daß es gewisse Gebiete $\mathfrak{R}^*$ gibt, die eine ausgezeichnete zweidimensionale Berandungsfläche $\mathfrak{C}$ besitzen: An keiner Stelle der (dreidimensionalen) Berandung kann der Absolutwert $|f(X, Z)|$ einer in $\mathfrak{R}^*$ regulären Funktion größer sein als das Maximum von $|f(X, Z)|$ *auf* $\mathfrak{C}$.

Es läßt sich daraus für diese Gebiete ein Analogon des Schwarzschen Lemma erschließen.

§ 1.

Über die Repräsentanten von Klassen von Bereichen.

Es läßt sich bekanntlich durch ein normiertes Funktionenpaar nicht jeder Bereich $\mathfrak{B}$ auf einen anderen $\mathfrak{B}^*$ abbilden. Die Gesamtheit derjenigen Bereiche, die durch (in bezug auf einen festen Punkt) normierte Funktionenpaare ineinander überführbar sind, fassen wir zu einer Klasse zusammen.

Wie in der Einleitung erwähnt wurde, besteht unsere Aufgabe darin, einen ausgezeichneten Repräsentanten in einer Klasse auszusondern und ein Verfahren anzugeben, zu einem Bereich dasjenige Funktionenpaar zu finden, welches die Abbildung von $\mathfrak{B}$ auf den Repräsentantenbereich leistet.

Wir beschränken unsere Untersuchung auf die einfach zusammenhängenden[14]), ganz im Endlichen gelegenen, von endlich vielen regulären dreidimensionalen Mannigfaltigkeiten berandeten schlichten Bereiche $\mathfrak{B}$, die den Punkt $0, 0$ im Innern enthalten. Wir führen gewisse Matrizen ein und werden uns der üblichen Symbole des Matrizenkalküls bedienen.

Es sei

$$(1) \quad \mathfrak{a}(X, Z) \equiv \begin{pmatrix} \sum \Omega_\nu(X, Z)\,\bar{\Omega}_\nu & \sum \Omega_\nu(X, Z)\,\bar{\Omega}'_{\nu X} & \sum \Omega_\nu(X, Z)\,\bar{\Omega}'_{\nu Z} \\ \sum \Omega_\nu \bar{\Omega}_\nu & \sum \Omega_\nu \bar{\Omega}'_{\nu X} & \sum \Omega_\nu \bar{\Omega}'_{\nu Z} \\ \sum \Omega'_{\nu X} \bar{\Omega}_\nu & \sum \Omega'_{\nu X} \bar{\Omega}'_{\nu X} & \sum \Omega'_{\nu X} \bar{\Omega}'_{\nu Z} \\ \sum \Omega'_{\nu Z} \bar{\Omega}_\nu & \sum \Omega'_{\nu Z} \bar{\Omega}'_{\nu X} & \sum \Omega'_{\nu Z} \bar{\Omega}'_{\nu Z} \end{pmatrix}$$

[14]) Unter einem einfach zusammenhängenden Bereich wird hier stets ein zusammenhängender Bereich verstanden, in dem jede ein-, zwei- oder dreidimensionale geschlossene Mannigfaltigkeit, die ganz in dessen Innerem liegt, sich in dem Bereich stetig auf einen Punkt zusammenziehen läßt.

eine Matrix, worin $\Omega_s(X, Z)$ $[s = 1, 2, 3, \ldots]$ das vollständige Orthogonalfunktionensystem zu $\mathfrak{B}$ bedeutet; es gelten ferner die Abkürzungen (3) und (4) der Einleitung und die Abkürzungen

$$(2) \qquad \mathfrak{a} = \mathfrak{a}(0,0),$$

$$(3) \qquad \mathsf{K}(X, Z; a, b) = \sum_{s=1}^{s=\infty} \Omega_s(X, Z)\,\overline{\Omega_s(a, b)}.$$

Die unteren Indizes bei den Matrizen sollen die fehlenden Kolonnen und Zeilen angeben, wobei die römischen Zahlen sich auf die fehlenden Kolonnen, die arabischen auf die fehlenden Zeilen beziehen. Nebeneinander und untereinander stehende deutsche Buchstaben, z. B. $(\mathfrak{a}, \mathfrak{b})$; $\begin{pmatrix} 0 & \mathfrak{b}' \\ \mathfrak{b} & \mathfrak{a} \end{pmatrix}$, bedeuten, daß die entsprechenden Matrizen formal nebeneinander bzw. untereinander aufgeschrieben sind. Der Strich über einer Größe bedeutet den Übergang zu der konjugiert Komplexen.

Es gilt nun der

Satz. *Die Funktionen*

$$(4) \qquad \begin{aligned} v(X, Z) &= -\frac{|\mathfrak{a}_3(X, Z)|}{\mathsf{K}(X, Z; 0, 0)} \, \frac{\mathsf{K}(0, 0; 0, 0)}{|\mathfrak{a}_1|} \\ w(X, Z) &= \frac{|\mathfrak{a}_4(X, Z)|}{\mathsf{K}(X, Z; 0, 0)} \, \frac{\mathsf{K}(0, 0; 0, 0)}{|\mathfrak{a}_1|} \end{aligned}$$

bilden ein normiertes Kovariantenpaar (vgl. Fußnote [11]), S. 431).

Der durch (4) *erzeugte Bildbereich von* $\mathfrak{B}$ *hat die Eigenschaft, daß er allein durch die Bereichklasse* (zu der $\mathfrak{B}$ gehört) *bestimmt ist.*

Diesen Bereich können wir somit als den Repräsentantenbereich der betreffenden Klasse verwenden.

Wir werden diese Repräsentantenbereiche im folgenden mit $\mathfrak{A}$ bezeichnen. (Die Repräsentantenbereiche brauchen natürlich im allgemeinen nicht schlicht zu sein.)

Beweis des Satzes. In der Arbeit H wurde gezeigt, daß es zu $\mathfrak{B}$ eine und nur eine in $\mathfrak{B}$ reguläre Funktion $u(X, Z)$ gibt, die, an Stelle von t eingesetzt, dem Integral

$$(5) \qquad \iiiint\limits_{\mathfrak{B}} t(X, Z)\,\overline{t(X, Z)}\, d\omega$$

einen Minimalwert erteilt, wobei zum Vergleich alle in $\mathfrak{B}$ regulären Funktionen mit der Nebenbedingung

$$(6) \qquad t(0,0) = 1$$

herangezogen werden. $d\omega$ bedeutet dabei das (vierdimensionale) Volumenelement.

$u(X, Z)$ wurde als die Minimalfunktion bezeichnet. $u(X, Z)$ ist bekanntlich

$$u(X, Z) = \frac{\mathsf{K}(X, Z; 0, 0)}{\mathsf{K}(0, 0; 0, 0)}.$$

Wir wollen zeigen, daß eine andere, ähnliche Variationsaufgabe eine und nur eine Lösung besitzt: Es soll eine Funktion $u^{(1)}(X, Z)$ gesucht sein, die ebenfalls dem Integral (5) den Minimalwert erteilt, wobei jetzt anstatt (6) die Nebenbedingungen

(7) $$t(0, 0) = 0, \quad \left[\frac{\partial t(X, Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = 1, \quad \left[\frac{\partial t(X, Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = 0$$

gelten sollen.

Es ist

(8) $$u^{(1)}(X, Z) = -\frac{|\mathfrak{b}, \mathfrak{a}(X, Z)|}{|\mathfrak{a}_1|},$$

wo $\mathfrak{b}$ die Matrix $\begin{Bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{Bmatrix}$ bedeutet.

Den Beweis dieser Behauptung führen wir Schritt für Schritt dem entsprechenden Beweis (S. 649—657) in der Arbeit H gleich. Wir betrachten die folgende (endliche) Extremumaufgabe:

Wir wollen in dem Ausdruck

(9) $$U_p(X, Z) = \sum_{s=1}^{s=p} A_s \Omega_s(X, Z)$$

die Konstanten A_s so bestimmen, daß derselbe, an Stelle von t in das Integral (5) eingesetzt, (5) zu einem Minimum macht, wobei zur Konkurrenz nur Ausdrücke von der Form (9), die die Nebenbedingungen (7) befriedigen, zugelassen werden. Bezeichnet man mit $\mathfrak{a}^{(p)}(X, Z)$ die Matrix (1), wobei jetzt die Summationen $\sum$, im Gegensatz zum früheren, $\sum_{s=1}^{s=p}$ bedeuten, so ist die Lösung dieser Aufgabe $u_p^{(1)}$ (wie die formale Rechnung zeigt)

(10) $$u_p^{(1)}(X, Z) = -\frac{|\mathfrak{b}, \mathfrak{a}^{(p)}(X, Z)|}{|\mathfrak{a}_1^{(p)}|}$$

gleich, und der Wert des Minimums $\lambda_1^{(p)}$

(11) $$\lambda_1^{(p)} = -\frac{\begin{vmatrix} 0 & \mathfrak{b}_1' \\ \mathfrak{b}_1 & \mathfrak{a}_1^{(p)} \end{vmatrix}}{|\mathfrak{a}_1^{(p)}|} = \frac{|\mathfrak{a}_{\mathrm{II}\,13}^{(p)}|}{|\mathfrak{a}_1^{(p)}|},$$

wo $\mathfrak{b}_1'$ üblicherweise die Transponierte zu $\mathfrak{b}_1$ bedeutet.

Wir wollen nun zeigen:

1. Es existiert

$$u^{(1)}(X, Z) = \lim_{p \to \infty} u_p^{(1)}(X, Z). \tag{12}$$

Die Konvergenz der einzelnen in der Matrix (1) auftretenden Summen $\sum_{s=1}^{s=\infty} \Omega_s(X, Z)\overline{\Omega}_s$, $\sum_{s=1}^{s=\infty} \Omega_s \overline{\Omega}'_{sX}$, ... folgt aus der Endlichkeit und der gleichmäßigen Konvergenz des Kernes $\sum_{s=1}^{s=\infty} |\Omega_s(a, b)|^2$ (vgl. Satz I und II der Arbeit H). $u^{(1)}(X, Z)$ stellt in jedem ganz im Innern von $\mathfrak{B}$ gelegenen Teilbereiche $\mathfrak{B}_m$ eine reguläre Funktion dar.

2. Es gibt keine Funktion $h^{(1)}(X, Z)$, die die Nebenbedingungen (7) erfüllt und dem Integral (5) einen kleineren Wert als $\lambda_1 = \lim_{p \to \infty} \lambda_1^{(p)} = \frac{|\mathfrak{a}_{\mathrm{II}\,13}|}{|\mathfrak{a}_1|}$ erteilt: Denn aus der Vollständigkeit des Systems $\Omega_s(X, Z)$ folgt, daß man zu jedem ε ein p bestimmen kann, derart, daß

$$\iiiint\limits_{\mathfrak{B}} h^{(1)}(X, Z)\, \overline{h^{(1)}(X, Z)}\, d\omega - \sum_{s=1}^{s=p} \Big| \iiiint\limits_{\mathfrak{B}} h^{(1)}(X, Z)\, \overline{\Omega_s(X, Z)}\, d\omega \Big|^2 \leq \varepsilon \tag{13}$$

wird. Bezeichnet man mit $h_p^{(1)}(X, Z)$

$$h_p^{(1)}(X, Z) = \sum_{s=1}^{s=p} \Omega_s(X, Z) \iiiint\limits_{\mathfrak{B}} h^{(1)}(X, Z)\, \overline{\Omega_s(X, Z)}\, d\omega, \tag{14}$$

so ist nach dem Hilfssatz 1 der Arbeit H (S. 649)

$$\begin{aligned}
&[h_p^{(1)}(X, Z)]_{\substack{X=0\\Z=0}} - [h^{(1)}(X, Z)]_{\substack{X=0\\Z=0}} = \varepsilon' && |\varepsilon'| \leq \sqrt{\frac{2\varepsilon}{\pi^2 \gamma_{ab}^4}},\\
&\left[\frac{\partial h_p^{(1)}(X, Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} - \left[\frac{\partial h^{(1)}(X, Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = \varepsilon'_X && |\varepsilon'_X| \leq \sqrt{\frac{6\varepsilon}{\pi^2 \gamma_{ab}^6}},\\
&\left[\frac{\partial h_p^{(1)}(X, Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} - \left[\frac{\partial h^{(1)}(X, Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = \varepsilon'_Z && |\varepsilon'_Z| \leq \sqrt{\frac{6\varepsilon}{\pi^2 \gamma_{ab}^6}}.
\end{aligned} \tag{15}$$

Berechnet man den Minimalwert M_p von (5), den ein Ausdruck von der Form (9) unter den Nebenbedingungen

$$t(X, Z)_{\substack{X=0\\Z=0}} = \varepsilon', \quad \left[\frac{\partial t(X, Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = 1 + \varepsilon'_X, \quad \left[\frac{\partial t(X, Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = \varepsilon'_Z \tag{16}$$

erteilt, so erhält man

$$M_p = - \frac{\begin{vmatrix} 0 & \bar{\varepsilon} \\ \varepsilon' & \mathfrak{a}_1^{(p)} \end{vmatrix}}{|\mathfrak{a}_1^{(p)}|}, \tag{17}$$

wobei $\mathfrak{e}$ die Matrix

$$\mathfrak{e} = \{\varepsilon',\ 1 + \varepsilon'_X,\ \varepsilon'_Z\}$$

und $\mathfrak{e}'$ die dazu Transponierte $\begin{Bmatrix} \varepsilon' \\ 1 + \varepsilon'_X \\ \varepsilon'_Z \end{Bmatrix}$ bedeutet. Es ist somit

$$(18)\qquad \iiiint_{\mathfrak{B}} h^{(1)}(X, Z)\,\overline{h^{(1)}(X, Z)}\,d\omega \geqq -\operatorname*{limes}_{p\to\infty} \frac{\begin{vmatrix} 0 & \bar{\mathfrak{e}} \\ \mathfrak{e}' & \mathfrak{a}_1^{(p)} \end{vmatrix}}{|\mathfrak{a}_1^{(p)}|}.$$

Da man ε', ε'_X, ε'_Z beliebig klein machen kann, folgt, daß:

$$(19)\qquad \iiiint_{\mathfrak{B}} h^{(1)}(X, Z)\,\overline{h^{(1)}(X, Z)}\,d\omega \geqq -\frac{\begin{vmatrix} 0 & \mathfrak{b}'_1 \\ \mathfrak{b}_1 & \mathfrak{a}_1 \end{vmatrix}}{|\mathfrak{a}_1|}$$

ist, w. z. b. w.

3. Es ist

$$(20)\qquad \iiiint_{\mathfrak{B}} u^{(1)}(X, Z)\,\overline{u^{(1)}(X, Z)}\,d\omega = -\frac{\begin{vmatrix} 0 & \mathfrak{b}'_1 \\ \mathfrak{b}_1 & \mathfrak{a}_1 \end{vmatrix}}{|\mathfrak{a}_1|}.$$

Denn in jedem (ganz im Innern von $\mathfrak{B}$ gelegenen) Teilbereiche $\mathfrak{B}_m$ ist bei genügend großem p

$$\left|\, |u^{(1)}(X, Z)|^2 - \left| \sum_{s=1}^{s=p} \Omega_s(X, Z) \frac{\begin{vmatrix} \mathfrak{l}_s \\ \mathfrak{a}_{12} \end{vmatrix}}{|\mathfrak{a}_1|} \right|^2 \right| \leqq \varepsilon,$$

wo $\mathfrak{l}_s$ die Matrix $\{\bar{\Omega}_s,\ \bar{\Omega}'_{sX},\ \bar{\Omega}'_{sZ}\}$ bedeutet, und

$$\iiiint_{\mathfrak{B}_m} |u^{(1)}(X, Z)|^2\,d\omega \leqq \iiiint_{\mathfrak{B}_m} \left| \sum_{s=1}^{s=p} \Omega_s(X, Z) \frac{\begin{vmatrix} \mathfrak{l}_s \\ \mathfrak{a}_{12} \end{vmatrix}}{|\mathfrak{a}_1|} \right|^2 d\omega + \varepsilon\,(\text{Volumen von } \mathfrak{B}_m) \leqq$$

$$(21)\qquad \leqq \iiiint_{\mathfrak{B}} \left| \sum_{s=1}^{s=p} \Omega_s(X, Z) \frac{\begin{vmatrix} \mathfrak{l}_s \\ \mathfrak{a}_{12} \end{vmatrix}}{|\mathfrak{a}_1|} \right|^2 d\omega + \varepsilon\,(\text{Volumen von } \mathfrak{B}) =$$

$$= \sum_{s=1}^{s=p} \frac{\left|\begin{vmatrix} \mathfrak{l}_s \\ \mathfrak{a}_{12} \end{vmatrix}\right|^2}{\big||\mathfrak{a}_1|\big|^2} + \varepsilon\,(\text{Volumen von } \mathfrak{B}) \leqq \frac{\sum_{s=1}^{s=\infty} \left|\begin{vmatrix} \mathfrak{l}_s \\ \mathfrak{a}_{12} \end{vmatrix}\right|^2}{\big||\mathfrak{a}_1|\big|^2} + \varepsilon\,(\text{Volumen von } \mathfrak{B}) =$$

$$= -\frac{\begin{vmatrix} 0 & \mathfrak{b}'_1 \\ \mathfrak{b}_1 & \mathfrak{a}_1 \end{vmatrix}}{|\mathfrak{a}_1|} + \varepsilon\,(\text{Volumen von } \mathfrak{B}).$$

Da man ε beliebig klein machen kann, folgt (20).

4. Da das System $\Omega_s(X, Z)$ in $\mathfrak{B}$ vollständig ist, ist für $u^{(1)}(X, Z)$ die Vollständigkeitsrelation erfüllt.

5. Es soll nunmehr der **Eindeutigkeitsbeweis** der Lösung der Variationsaufgabe geführt werden. Sei $G(X, Z)$ eine Funktion, die den Nebenbedingungen (7) genügt und für die

$$\iiiint_{\mathfrak{B}} G(X, Z)\overline{G(X, Z)}\, d\omega = \frac{|\mathfrak{a}_{\mathrm{II}\,13}|}{|\mathfrak{a}_1|} \tag{22}$$

ist. Wir werden zeigen, daß

$$G(X, Z) = -\frac{|\mathfrak{b}, \mathfrak{a}(X, Z)|}{|\mathfrak{a}_1|} \tag{23}$$

sein muß. Wir führen noch die folgenden Abkürzungen ein:

$$\begin{gathered} \mathfrak{b}(X, Z) = \left\{ \sum \Omega''_{sX^2}(X, Z)\overline{\Omega_s},\quad \sum \Omega''_{sX^2}(X, Z)\overline{\Omega}'_{sX},\quad \sum \Omega''_{sX^2}(X, Z)\overline{\Omega}'_{sZ} \right\}, \\ \mathfrak{b} = \mathfrak{b}(0, 0),\quad \Omega''_{sX^2}(X, Z) = \frac{\partial^2 \Omega_s(X, Z)}{\partial X^2},\quad \Omega''_{sX^2} = [\Omega''_{sX^2}(X, Z)]_{\substack{X=0\\ Z=0}}, \end{gathered} \tag{24}$$

$$\sum = \sum_{s=1}^{s=\infty}.$$

Da $\Omega_s(X, Z)$ in $\mathfrak{B}$ vollständig ist, folgt, daß man zu jedem ε ein p derart finden kann, daß

$$\iiiint_{\mathfrak{B}} G(X, Z)\overline{G(X, Z)}\, d\omega - \sum_{s=1}^{s=p} |g_s|^2 \leqq \varepsilon, \tag{25}$$

$$g_s = \iiiint_{\mathfrak{B}} G(X, Z)\overline{\Omega_s(X, Z)}\, d\omega.$$

Betrachten wir nun die Funktion

$$G_p = \sum_{s=1}^{s=p} g_s \Omega_s(X, Z). \tag{26}$$

Nach dem Hilfssatz I der Arbeit H folgt, daß G_p die Nebenbedingungen (16) erfüllt. Man kann ferner die zweite Ableitung von G_p im Koordinatenanfangspunkte in der Form

$$\left[\frac{\partial^2 G_p(X, Z)}{\partial X^2}\right]_{\substack{X=0\\ Z=0}} = \sum_{s=1}^{s=p} g_s \left[\frac{\partial^2 \Omega_s(X, Z)}{\partial X^2}\right]_{\substack{X=0\\ Z=0}} = -\frac{\begin{vmatrix} 0 & \mathfrak{b}^{(p)} \\ \mathfrak{c}' & \mathfrak{a}_1^{(p)} \end{vmatrix}}{|\mathfrak{a}_1^{(p)}|} + R_p^{(2,0)} \equiv N_p + R_p^{(2,0)} \tag{27}$$

darstellen. Wir wollen zeigen, daß

$$\operatorname*{limes}_{p \to \infty} R_p^{(2,0)} = 0 \tag{28}$$

ist. Um dies einzusehen, betrachten wir das Minimum des Integrals (5), wobei an Stelle von t ein Ausdruck von der Form (9) zu nehmen ist, der sowohl die Nebenbedingungen (16) wie auch die Bedingung

$$\left[\frac{\partial^2 \left(\sum_{s=1}^{s=p} A_s \Omega_s(X, Z)\right)}{\partial X^2}\right]_{\substack{X=0\\ Z=0}} = N_p + R_p^{(2,0)} \tag{29}$$

erfüllen soll. Wie eine formale Rechnung zeigt, hat dieses Minimum den Wert

$$(30)\qquad \frac{\begin{vmatrix} 0 & \bar{\mathfrak{e}} & \overline{N_p} \\ \mathfrak{e}' & \mathfrak{a}_1^{(p)} & \overline{\mathfrak{b}^{(p)\prime}} \\ N_p & \mathfrak{b}^{(p)} & \sum\limits_{s=1}^{s=p} \Omega''_{sX^s}\overline{\Omega}''_{sX^s} \end{vmatrix} + |R_p^{(2,0)}|^2\,|\mathfrak{a}_1^{(p)}|}{\begin{vmatrix} \mathfrak{a}_1^{(p)} & \overline{\mathfrak{b}^{(p)\prime}} \\ \mathfrak{b}^{(p)} & \sum\limits_{s=1}^{s=p} \Omega''_{sX^s}\overline{\Omega}''_{sX^s} \end{vmatrix}} \equiv S_1^{(p)} + |R_p^{(2,0)}|^2\, S_2^{(p)}.$$

Da $G_p(X, Z)$ eine zulässige Konkurrenzfunktion ist, gilt:

$$(31)\qquad \sum_{s=1}^{s=p} |g_s|^2 \geqq S_1^{(p)} + |R_p^{(2,0)}|^2\, S_2^{(p)}.$$

Geht nun p gegen ∞, so ist (wie eine formale Rechnung zeigt)

$$(32)\qquad \operatorname*{limes}_{p\to\infty} S_1^{(p)} = -\operatorname*{limes}_{p\to\infty} \frac{\begin{vmatrix} 0 & \mathfrak{b}_1' \\ \mathfrak{b}_1 & \mathfrak{a}_1^{(p)} \end{vmatrix}}{|\mathfrak{a}_1^{(p)}|} = -\frac{\begin{vmatrix} 0 & \mathfrak{b}_1' \\ \mathfrak{b}_1 & \mathfrak{a}_1 \end{vmatrix}}{|\mathfrak{a}_1|}.$$

$S_2^{(p)}$ konvergiert gegen eine positive, von 0 verschiedene Zahl: Suchen wir nämlich das Minimum von (5) bei den Nebenbedingungen

$$(33)\quad t(0,0)=0,\ \left[\frac{\partial t(X,Z)}{\partial X}\right]_{\substack{X=0\\Z=0}}=0,\ \left[\frac{\partial t(X,Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}}=0,\ \left[\frac{\partial^2 t(X,Z)}{\partial X^2}\right]_{\substack{X=0\\Z=0}}=1,$$

so erhalten wir $\operatorname*{limes}_{p\to\infty} S_2^{(p)}$ als den Wert des offenbar positiven Minimums. (Nach dem Hilfssatz I der Arbeit H muß nämlich $\operatorname*{limes}_{p\to\infty} S_2^{(p)} \geqq \frac{12}{\pi^2 \gamma_{00}^8}$ sein, wo γ_{00} den Abstand des Punktes 0,0 von der Berandung bedeutet.) Da nach (22) und (25)

$$\operatorname*{limes}_{p\to\infty} \sum_{s=1}^{s=p} |g_s|^2 = -\frac{\begin{vmatrix} 0 & \mathfrak{b}_1' \\ \mathfrak{b}_1 & \mathfrak{a}_1 \end{vmatrix}}{|\mathfrak{a}_1|}$$

und $\operatorname*{limes}_{p\to\infty} S_2^{(p)}$ größer als 0 ist, so folgt aus (30) und (32), daß

$$(28)\qquad \operatorname*{limes}_{p\to\infty} R_p^{(2,0)} = 0$$

ist, w. z. b. w.

Genau auf dieselbe Weise läßt sich für jedes ν, μ zeigen, daß

$$(34)\quad \operatorname*{limes}_{p\to\infty} \left[\frac{\partial^{\nu+\mu} G_p(X,Z)}{\partial X^\nu \partial Z^\mu}\right]_{\substack{X=0\\Z=0}} = -\operatorname*{limes}_{p\to\infty} \left[\frac{\partial^{\nu+\mu}}{\partial X^\nu \partial Z^\mu}\left(\frac{|\mathfrak{b}, \mathfrak{a}^{(p)}(X,Z)|}{|\mathfrak{a}_1^{(p)}|}\right)\right]_{\substack{X=0\\Z=0}}$$

wird. Nach dem Satz II der Arbeit H und Punkt 1 dieses Beweises konvergieren in der Umgebung des Punktes 0, 0

(35) $$\sum_{s=1}^{s=p} g_s\,\Omega_s(X, Z) \quad \text{gegen} \quad G(X, Z)$$

(36) $$-\frac{|\mathfrak{b}, \mathfrak{a}^{(p)}(X, Z)|}{|\mathfrak{a}_1^{(p)}|} \quad \text{gegen} \quad -\frac{|\mathfrak{b}, \mathfrak{a}(X, Z)|}{|\mathfrak{a}_1|}.$$

Nach (34) konvergieren ferner in der Umgebung des Punktes 0, 0 alle Ableitungen von $\sum_{s=1}^{s=p} g_s\,\Omega_s(X, Z)$ gegen die entsprechenden Ableitungen von

$$-\frac{|\mathfrak{b}, \mathfrak{a}^{(p)}(X, Z)|}{|\mathfrak{a}_1^{(p)}|}.$$

Da $u^{(1)}(X, Z)$ und $G(X, Z)$ in der Umgebung des Punktes 0, 0 regulär analytisch sind, folgt daraus (23), w. z. b. w.

Auf gleichem Wege wie die Existenz und die Eindeutigkeit der Lösung $u^{(1)}(X, Z)$ bewiesen wurde, läßt sich zeigen, daß auch die folgende Variationsaufgabe eine und nur eine Lösung besitzt: Es ist diejenige Funktion gesucht, die (5) den Minimalwert erteilt, wobei als Nebenbedingungen die Relationen

(37) $$t(0, 0) = 0, \quad \left[\frac{\partial t(X, Z)}{\partial X}\right]_{\substack{X=0\\Z=0}} = 0, \quad \left[\frac{\partial t(X, Z)}{\partial Z}\right]_{\substack{X=0\\Z=0}} = 1$$

zu nehmen sind. Die Lösung dieser zweiten Aufgabe sei mit $u^{(2)}(X, Z)$ bezeichnet.

Wir gehen nunmehr zu dem zweiten Schritt des Beweises über: Seien $\mathfrak{B}$ und $\mathfrak{B}^*$ zwei Bereiche des XZ- bzw. X^*Z^*-Raumes, die zu derselben Klasse gehören, d. h. es gebe ein normiertes Funktionenpaar $X^*(X, Z)$, $Z^*(X, Z)$ das $\mathfrak{B}$ auf $\mathfrak{B}^*$ abbildet.

Verpflanzt man die Minimalfunktion $u(X, Z)$ in den X^*Z^*-Raum und multipliziert man mit der Funktionaldeterminante $\left(\text{d. h. bildet man } u[X(X^*, Z^*), Z(X^*, Z^*)]\,\frac{\partial(X, Z)}{\partial(X^*, Z^*)}\right)$, so ist dieser Ausdruck der Minimalfunktion $u^*(X^*, Z^*)$ von $\mathfrak{B}^*$ gleich. Denn:

1. Das Funktionselement um den Koordinatenanfangspunkt von

$$\begin{aligned} &u[X(X^*, Z^*), Z(X^*, Z^*)]\,\frac{\partial(X, Z)}{\partial(X^*, Z^*)} \\ &\equiv (1 + \alpha_{10} X^* + \alpha_{01} Z^* + \ldots)(1 + \beta_{10} X^* + \beta_{01} Z^* + \ldots) \\ &\equiv 1 + A_{10} X^* + A_{01} Z^* + \ldots\ . \end{aligned}$$

hat an der Stelle $X = 0$, $Z = 0$ den Wert 1.

2. Die Funktion $u[X(X^*, Z^*), Z(X^*, Z^*)] \frac{\partial(X, Z)}{\partial(X^*, Z^*)}$ erteilt dem Integral $\iiiint_{\mathfrak{B}^*} |t(X^*, Z^*)|^2 d\omega_{X^*Z^*}$ den Minimalwert $\lambda^* = \lambda$ [15]). Denn es ist:

$$(38)\quad \iiiint_{\mathfrak{B}^*} \left| u[X(X^*, Z^*), Z(X^*, Z^*)] \frac{\partial(X, Z)}{\partial(X^*, Z^*)} \right|^2 d\omega_{X^*Z^*} = \iiiint_{\mathfrak{B}} |u(X, Z)|^2 d\omega_{XZ}.$$

Nach dem Satz III der Arbeit H gibt es zu jedem Bereich nur eine Minimalfunktion $u^*(X^* Z^*)$. Mit anderen Worten: es ist

$$(39)\quad \frac{u[X(X^*, Z^*), Z(X^*, Z^*)]}{u^*(X^*, Z^*)} \frac{\partial(X, Z)}{\partial(X^*, Z^*)} = 1.$$

Das Funktionselement von

$$(40)\quad u^{(1)}[X(X^*, Z^*), Z(X^*, Z^*)] \frac{\partial(X, Z)}{\partial(X^*, Z^*)}$$
$$= [X^* + \alpha'_{20} X^{*2} + \alpha'_{11} X^* Z^* + \ldots][1 + \beta_{10} X^* + \beta_{01} Z^* + \ldots]$$
$$= [X^* + B_{20} X^{*2} + B_{11} X^* Z^* + \ldots]$$

um den Punkt $X^* = 0$, $Z^* = 0$ befriedigt die Nebenbedingungen (7). Auf Grund der früher bewiesenen Eindeutigkeit der Lösung der entsprechenden Variationsaufgabe gibt es eine und nur eine Funktion $u^{(1)*}(X^*, Z^*)$, die dem Integral (5) unter den Nebenbedingungen (7) den Minimalwert erteilt. Auf gleiche Weise wie früher schließen wir, daß

$$(41)\quad \frac{u^{(1)}[X(X^*, Z^*), Z(X^*, Z^*)]}{u^{(1)*}(X^*, Z^*)} \frac{\partial(X, Z)}{\partial(X^*, Z^*)} = 1$$

ist. Aus (39) und (41) folgt, daß für zwei Bereiche derselben Klasse $\mathfrak{B}$ und $\mathfrak{B}^*$

$$(42)\quad \frac{u^{(1)*}[X^*, Z^*]}{u^*[X^*, Z^*]} = \frac{u^{(1)}[X(X^*, Z^*), Z(X^*, Z^*)]}{u[X(X^*, Z^*), Z(X^*, Z^*)]} = \frac{u^{(1)}(X, Z)}{u(X, Z)}$$

ist. Bezeichnen wir mit $v^*(X^*, Z^*)$ den in (42) links stehenden, mit $v(X, Z)$ den rechts stehenden Ausdruck, so ist $v(X, Z)$ eine Kovariante, d. h. es ist

$$(42^*)\quad v(X, Z) = v^*(X^*, Z^*).$$

Eine zweite Kovariante erhalten wir in $w(X, Z) = \frac{u^{(2)}(X, Z)}{u(X, Z)}$.

$v(X, Z)$ und $w(X, Z)$ bilden ein normiertes Funktionenpaar. Die Funktionaldeterminante $\frac{\partial(v, w)}{\partial(X, Z)}$ ist bestimmt nicht identisch Null. Wir haben somit eine Abbildung auf einen Bereich gewonnen, der von der Wahl des Ausgangsgebietes $\mathfrak{B}$ unabhängig ist. Durch v und w wird somit der Bereich $\mathfrak{B}$ auf den Repräsentantenbereich $\mathfrak{A}$ abgebildet.

---

[15]) λ ist eine Invariante für alle Bereiche, die zu derselben Klasse gehören.

Ein Repräsentantenbereich $\mathfrak{A}$ ist dadurch ausgezeichnet, daß das durch (4) gegebene, zu $\mathfrak{A}$ gehörige Funktionenpaar sich auf $v = X$, $w = Z$ reduziert.

Korollar. *Ein Reinhardtscher Kreisbereich ist ein Repräsentantenbereich*, was man z. B. durch die unmittelbare Ausrechnung der durch (4) gegebenen Funktionen ersehen kann.

Gehört somit der Reinhardtsche Kreisbereich $\mathfrak{K}$ zur Klasse, so wird (4) das Abbildungsfunktionenpaar liefern, das $\mathfrak{B}$ auf $\mathfrak{K}$ abbildet. Die Gestalt von $\mathfrak{K}$ können wir in diesem Falle mit Hilfe der im § 2 der Arbeit H angegebenen Formel feststellen.

§ 2.
Über die Approximation der Funktionen von zwei komplexen Veränderlichen durch Polynome.

In der Arbeit „Zwei Sätze über Funktionen von zwei komplexen Veränderlichen“ wurde mit Hilfe des Auswahlprinzips die Existenz eines vollständigen Orthogonalfunktionensystems zu jedem zusammenhängenden Bereich $\mathfrak{B}$ nachgewiesen. Es entsteht das Problem, zu einem gegebenen Bereich ein konstruktives Verfahren anzugeben, um ein derartiges System zu bilden.

Ein naturgemäßer Weg dazu besteht im Falle der einfach zusammenhängenden Bereiche darin, daß man den Rungeschen Satz über die Approximation von analytischen Funktionen durch Polynome[16]) auf den Fall von zwei komplexen Veränderlichen überträgt. Aus diesem Satz folgt dann mühelos die Tatsache, daß das durch die Orthogonalisierung von

$$1, X, Z, X^2, XZ, Z^2, \ldots \tag{1}$$

erzeugte System in dem betreffenden Bereiche vollständig ist.

Herr Almar hat in seiner Dissertation „Sur quelques problèmes de la théorie des fonctions de deux variables complexes“ [17]) diesen Satz für gewisse Klassen von Bereichen nachgewiesen[18]) und die (unbewiesene)

[16]) C. Runge, Zur Theorie der eindeutigen analytischen Funktionen, Acta Mathematica 6 (1884), S. 229.

[17]) Arkiv för matematik, astronomi och fysik 17 (1922/23), Nr. 7, S. 1–70.

[18]) Herr Almar stützt sich dabei auf eine Verallgemeinerung der Mittag-Lefflerschen Darstellung einer Funktion: Ist $\sum_{(nm)=0}^{(nm)=\infty} c_{nm} X^n Z^m$ das Funktionselement der Funktion $f(X, Z)$ um den Koordinatenanfangspunkt, so läßt sich — wie Herr Almar zeigt — $f(X, Z)$ in gewissen Bereichen in der Form

$$\operatorname*{limes}_{\alpha \to 0} \sum_{(nm)=0}^{(nm)=\infty} \frac{c_{nm}}{\Gamma[1 + \alpha(m+n)]} X^n Z^m$$

darstellen.

Vermutung ausgesprochen, daß dieser Satz in Gebieten, die Herr Almar als quasikonvex bezeichnet, richtig ist.

Im folgenden soll ein Beweis des verallgemeinerten Rungeschen Satzes geführt werden, der eine Integraldarstellung der Funktion benutzt und der dem Rungeschen Beweis nachgebildet ist. Der Beweis bezieht sich ebenfalls nur auf gewisse Klassen von Bereichen. (Darunter befinden sich auch Bereiche, die von dem Almarschen Beweis nicht erfaßt wurden.)

Die Integraldarstellung (6) der Funktion, die in gewissen (weiter näher definierten) Gebieten gilt, erlaubt aus dem Verlauf der Funktion auf gewissen (zweidimensionalen) Flächen auf den Wertevorrat der Funktion in dem (vierdimensionalen) Gebiete zu schließen. Anschließend an den Beweis werden wir aus dieser Integraldarstellung einige einfache Folgerungen ziehen.

Wir führen den Beweis für den einfachsten Fall durch, doch läßt sich — wie in der Fußnote[19]) hingewiesen wird — derselbe bedeutend verallgemeinern.

Seien uns zwei analytische Ebenen etwa die Ebene

$$Z = \text{konst.} = c_1 \tag{2}$$

und

$$X + mZ = \text{konst.} = c_2 \tag{3}$$

gegeben. Seien ferner B und B' zwei ebene, einfach zusammenhängende, konvexe Bereiche, die von den regulären Kurven C bzw. C' berandet sind. B liege in der Ebene (2), B' in (3). Es bedeute $x^{(i)}, c_1$ den laufenden Punkt von B; $x^{(b)}, c_1$ von C; $(c_2 - mz^{(i)}), z^{(i)}$ von B'; $(c_2 - mz^{(b)}), z^{(b)}$ von C'. $\mathfrak{P}$ bedeute einen Bereich des P_4, dessen Inneres aus der Gesamtheit aller Punkte X, Z besteht, die die Relationen

$$X + mZ = x^{(i)} + mc_2, \tag{4}$$

$$Z = z^{(i)} \tag{4*}$$

befriedigen. $\mathfrak{C}$ bedeute eine (zweidimensionale) Fläche, die aus der Gesamtheit aller Punkte X, Z besteht, die die Gleichungen

$$X + mZ = x^{(b)} + mc_2, \tag{5}$$

$$Z = z^{(b)} \tag{5*}$$

befriedigen. Wir verstehen unter $\mathfrak{R}$ einen solchen einfach zusammenhängenden Teilbereich von $\mathfrak{P}$, dessen (dreidimensionale) Berandungsmannigfaltigkeit $\mathfrak{C}$ enthält. Es gilt nun der

Satz. *Eine im abgeschlossenen Bereiche $\mathfrak{R}$ reguläre Funktion $f(X, Z)$ läßt sich in jedem ganz im Innern von R gelegenen Teilbereiche $\mathfrak{R}'$ von $\mathfrak{R}$ gleichmäßig durch Polynome approximieren.*

Beweis. Nach Poincaré können wir $f(X, Z)$ in $\mathfrak{P}$ in der Form

$$f(X, Z) = -\frac{1}{4\pi^2}\iint\limits_{\mathfrak{C}} \frac{f(x, z)\, d\tau}{[Z - z][(X - x) + m(Z - z)]} \tag{6}$$

darstellen[19]. $d\tau$ bedeutet darin $\frac{\partial(x, z)}{\partial(N, M)}\, dN\, dM$, wobei $x = x(N, M)$, $z = z(N, M)$ die Parameterdarstellung der Fläche $\mathfrak{C}$ ist.

Wir ersetzen in (6) das Integral durch eine endliche Summe

$$\sum_k \sum \frac{f(x_k, z_k)\, \Delta_k}{[Z - z_k][(X - x_k) + m(Z - z_k)]}. \tag{7}$$

Man kann bei einer genügend feinen Unterteilung der Fläche $\mathfrak{C}$ erreichen, daß der Unterschied zwischen (6) und (7) gleichmäßig in jedem ganz im Innern von $\mathfrak{P}$ gelegenen Teilbereich $\mathfrak{P}'$ kleiner als eine vorgegebene kleine Zahl ε wird.

Wir gehen nunmehr dazu über, den einzelnen Summanden

$$\frac{f(x_k, z_k)\, \Delta_k}{[Z - z_k][(X - x_k) + m(Z - z_k)]}$$

durch ein Polynom (in X, Z) in $\mathfrak{P}'$ zu approximieren.

1. Wir wollen zunächst uns überlegen, daß $(X - x_k) + m(Z - z_k)$ (wo x_k, z_k ein Punkt von $\mathfrak{C}$ ist) nirgends im Innern von $\mathfrak{P}$ verschwindet. Jeder

---

[19]) Vgl. etwa Picard und Simart, Théorie des fonctions algébriques de deux variables indépendantes, Paris 1897, I. Bd., III. Kap., § 5, S. 55.

Die Formel (6) ist ein Spezialfall der Darstellungen

$$f(X, Z) = -\frac{1}{4\pi^2}\left[\frac{\partial R(X,Z)}{\partial Z}\frac{\partial Q(X,Z)}{\partial X} - \frac{\partial R(X,Z)}{\partial X}\frac{\partial Q(X,Z)}{\partial Z}\right]\iint\limits_{\mathfrak{C}} \frac{f(x,z)\, d\tau}{Q[(X-x),(Z-z)]\, R[(X-x),(Z-} \tag{6*}$$

wo Q und P Polynome sind, die im Koordinatenanfangspunkt eine gewöhnliche Nullstelle haben und sonst nirgends im Gebiete verschwinden, bzw. von

$$f(X, Z) = -\frac{1}{4\pi^2}\sqrt{(A''_{XZ})^2 - A''_{X^2} A''_{Z^2}}\iint\limits_{\mathfrak{C}} \frac{f(x, z)\, d\tau}{A[(X - x), (Z - z)]}, \tag{6**}$$

wo A ein Polynom zweiten Grades ist.

Mit Hilfe der Darstellungen (6*) und (6**) läßt sich der Rungesche Satz für eine weitere Reihe von Gebieten beweisen.

Die Schwierigkeit, die in diesem Beweis auftritt, besteht eigentlich im folgenden: Es ist im allgemeinen nicht möglich zu entscheiden, ob auf der Berandung eines vorgegebenen Bereiches $\mathfrak{B}$ eine Fläche $\mathfrak{C}$ existiert, zu der sich ein Polynomenpaar Q und R konstruieren läßt, mit dessen Hilfe man jede in $\mathfrak{B}$ reguläre Funktion in der Form (6*) darstellen kann, und wo die im Nenner stehenden Polynome für kein $(X, Z) = (x^{(0)}, z^{(0)})$ verschwinden.

Die Konvexität und die Regularität von C und C' sind bei diesen Überlegungen nicht wesentlich.

Punkt von $\mathfrak{C}$ genügt den Gleichungen (5) und (5*). Die Ebene $(X - x_k) + m(Z - z_k) = 0$ kann man ebenfalls in der Form

$$X + mZ = x^{(b)} + mc_2 \tag{8}$$

schreiben. Da zwei parallele Ebenen im Endlichen keinen gemeinsamen Punkt haben können, so hat keine der Ebenen (8) einen gemeinsamen Punkt mit der Ebene (4), also keinen Punkt, der im Innern von $\mathfrak{P}$ liegt. Ebenso zeigen wir, daß $Z - z_k$ nirgends in $\mathfrak{P}'$ verschwindet.

2. Wir wollen nunmehr $\frac{1}{m(Z - z_k) + (X - x_k)}$ durch ein Polynom approximieren. Wir legen durch den Punkt $x = mz_k + x_k - mc_2$ eine Tangente an B. Durch jeden Punkt dieser Tangente legen wir eine analytische Ebene

$$X + mZ = \text{konst.} \tag{9}$$

Die Gesamtheit dieser Ebenen liefert einen dreidimensionalen Raum, den wir mit T bezeichnen. Eine ähnliche Überlegung wie die unter 1. durchgeführte, zeigt, daß T keinen gemeinsamen Punkt mit dem Teilbereich $\mathfrak{P}'$ hat. Im Punkte x_k, z_k errichten wir zu T eine Normale und legen eine Hyperkugel, die T in x_k, z_k tangiert und deren Mittelpunkt somit auf der Normalen liegt. Der Radius dieser Hyperkugel soll endlich sein, jedoch so groß, daß $\mathfrak{P}'$ ganz im Innern dieser Hyperkugel liegt. Da

$$\frac{1}{(X - x_k) + m(Z - z_k)} \tag{10}$$

im Innern dieser Hyperkugel regulär ist, läßt sich (10) in der Hyperkugel in eine Potenzreihe entwickeln[20]). Nehmen wir einen genügend langen Abschnitt der Reihe, so wird der Unterschied zwischen dem Abschnitt und (10) gleichmäßig in $\mathfrak{P}'$ kleiner wie eine vorgegebene Größe ε. Wir haben also (10) durch ein Polynom approximiert; ebenso läßt sich $\frac{1}{Z - z_k}$ in $\mathfrak{P}'$ approximieren, woraus ohne Mühe unser Satz folgt[19]).

Es sollen einige weitere Folgerungen aus der Darstellung (6) gezogen werden, wobei wir uns auf den Fall $m = 0$ beschränken.

Der Absolutwert jeder in $\mathfrak{R}$ (somit in $\mathfrak{P}$) regulären Funktion $f(X, Z)$ wird an keiner Stelle in $\mathfrak{R}$ größer als das Maximum von $|f(X, Z)|$ a u f $\mathfrak{C}$. Denn: benutzt man die auf der Seite 443 eingeführten Bezeichnungen, so ist nach den Sätzen der (ebenen) Funktionentheorie:

$$|f(x^{(i)}, z^{(i)})| \leq \text{Max. von } |f(x^{(b)}, z^{(i)})| \leq \text{Max. von } |f(x^{(b)}, z^{(b)}|.$$

[20]) Vgl. dazu F. Hartogs, Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlicher, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten. Math. Annalen 62 (1905), S. 9 oder W. F. Osgood, Lehrbuch der Funktionentheorie, Bd. II, I. Lieferung, B. G. Teubner, Leipzig und Berlin 1924, S. 181.

Hat die Berandung von $\mathfrak{P}$ nur die Fläche $\mathfrak{C}$ mit $\mathfrak{R}$ gemeinsam, so wird $|f(X, Z)|$ — $\left[\text{sobald weder } \frac{\partial f}{\partial X} \text{ noch } \frac{\partial f}{\partial Z} \text{ identisch verschwinden}\right]$ an jeder Stelle in $\mathfrak{R}$ kleiner als das Maximum von $|f(X, Z)|$ auf $\mathfrak{C}$ sein. Wir werden solche Bereiche mit $\mathfrak{R}^*$, die Fläche $\mathfrak{C}$ als die Maximumfläche von $\mathfrak{R}^*$ bezeichnen.

Durch eine Abbildung von $\mathfrak{R}^*$ durch ein Paar von Funktionen von zwei komplexen Veränderlichen gelangen wir zu einem neuen Bereich, der ebenfalls eine Maximumfläche besitzt.

Werden zwei einfach zusammenhängende Bereiche, die Maximumflächen besitzen, aufeinander abgebildet, so geht die Maximumfläche des einen Bereiches in die Maximumfläche des anderen über.

Es läßt sich in üblicher Weise für die Bereiche $\mathfrak{R}$ (wie auch für die durch die analytische Abbildung aus ihnen erzeugte) das Schwarzsche Lemma erschließen:

Seien $f(X, Z)$ und $\varphi(X, Z)$ zwei in $\mathfrak{R}$ reguläre Funktionen, es sei ferner $\frac{f(X, Z)}{\varphi(X, Z)}$ in $\mathfrak{R}$ ebenfalls regulär.

Sei M das Maximum von $|f(X, Z)|$, μ das Minimum von $|\varphi(X, Z)|$ auf $\mathfrak{C}$. Es ist dann in jedem Punkte von $\mathfrak{R}$:

$$|f(X, Z)| \leqq \frac{M}{\mu} |\varphi(X, Z)|. \tag{11}$$

(Eingegangen am 10. 1. 1929.)

Die Methode der Polarfunktionen und Konfigurationskonstanten höherer Ordnung im Gebiete der Randwertaufgaben der Potentialtheorie[1].

Von

Ernst Richard Neumann in Marburg.

Der Nachweis für die Anwendbarkeit der Methode des arithmetischen Mittels von Carl Neumann zur Lösung der ersten Randwertaufgabe der Potentialtheorie war bekanntlich zunächst lange Jahre hindurch auf Gebiete beschränkt, die von *konvexen* Flächen σ begrenzt werden[2]). Bei diesem Beweise handelt es sich u. a. darum, zu zeigen, daß gewisse nach bestimmtem Verfahren sukzessive auf σ gebildete Funktionen $f', f'', f''', \ldots$ sich einer Konstanten nähern. Diesen Nachweis führte C. Neumann in der Weise, daß er die *Schwankungen* $\Delta(f^{(n)})$ dieser sukzessiven Funktionen $f^{(n)}$ betrachtete und für sie bei konvexer Fläche σ die Beziehung bewies:

$$\text{(N)} \qquad \Delta(f^{(n)}) \leqq \Delta(f^{(n-1)}) \cdot c_g, \qquad \text{also} \qquad \Delta(f^{(n)}) \leqq \Delta \cdot c_g^n,$$

wo $\Delta = \Delta(f)$ die Schwankung der vorgegebenen Randwertfunktion f bedeutet, und c_g eine der Fläche σ eigentümliche, also nur von den geometrischen Verhältnissen abhängige[3]) positive Konstante, die sogenannte

---

[1]) Ausführlichere Darstellung eines am 15. September 1925 auf der Mathematiker-Tagung in Danzig gehaltenen Vortrages. Jahresbericht der Dtsch. Math.-Vereinigung **34**, 2. Abt., S. *138–143*.

[2]) Der Einfachheit halber werde ich mich auf *räumliche* Gebiete beschränken. — Zur Bezeichnungsweise sei bemerkt, daß Innen- und Außengebiet von σ mit $\mathfrak{J}$ bzw. $\mathfrak{A}$ bezeichnet werden, Punkte innerhalb dieser Gebiete mit $i\,(\mathfrak{i})$ bzw. $a\,(\mathfrak{a})$, Punkte auf σ selber mit $s\,(\mathfrak{s})$, oder aber sie werden, wenn nötig, je nachdem man sie als Randpunkte von $\mathfrak{J}$ oder $\mathfrak{A}$ ansieht, noch genauer durch die Doppelindizes is bzw. as angedeutet. — Wenn Formeln gültig sind im Gebiete $\mathfrak{J}$ mit Einschluß der inneren Randpunkte, so sprechen wir von ihrer Gültigkeit in $\mathfrak{J}+\sigma$ (und analog beim Außengebiet).

[3]) Das soll nach dem Vorgange von C. Neumann durch die *Marke* g auch äußerlich angedeutet werden — die also nichts mit den später auftretenden Randwerten g, zu tun hat.

Konfigurationskonstante ist, die bei konvexen Flächen einen echtgebrochenen Wert besitzt, womit dann augenscheinlich der gewünschte Beweis *für konvexe Gebiete* im wesentlichen geführt ist[4]).

Erst Poincaré gab den Anstoß zur Erweiterung des Gültigkeitsbereiches der Neumannschen Methode, indem er nach dem Vorbilde von H. A. Schwarz eigentümliche Integrale J_n (oder $J^{(n)}$, wie wir schreiben werden) in den Kreis seiner grundlegenden Betrachtungen zog[5]), die hinüberleiteten zur Theorie der Integralgleichungen, deren man sich heute zu bedienen pflegt, um die Allgemeingültigkeit der Methode des arithmetischen Mittels zu beweisen. — Zumal hier aber Integralgleichungen mit unstetigen und unsymmetrischen Kernen auftreten, bedarf es bei diesem Wege eines ziemlich weiten Ausholens, und es dürfte daher auch heute noch ein direkter Beweis für die Konvergenz der Methode oder eine Vervollkommnung der älteren Beweise, wie diese Arbeit sie bringen will, nicht wertlos sein, zumal zu hoffen ist, daß diese Vervollkommnungen vielleicht auch wieder der Integralgleichungstheorie zugute kommen werden.

Mit Hilfe der von mir eingeführten „Polarfunktionen" habe ich bereits früher[6]), wenigstens für die eine in den Poincaréschen Betrachtungen enthaltene Tatsache, einen verhältnismäßig einfachen Beweis geliefert, nämlich (beim Innengebiet) für die Formel:

$$\text{(P)} \qquad \operatorname{abs}(W_i^{(n)} - W_i^{(n+1)}) \leq A_g \sqrt{J^{(2n-2)}},$$

in der die $W^{(n)}$ die aufeinanderfolgenden Neumannschen Potentiale bedeuten und A_g wieder eine solche nur von den geometrischen Verhältnissen abhängige Konstante ist. — Die Bedeutung dieser Formel besteht darin, daß sie gestattet — und darin möchte ich das wesentliche Prinzip der Poincaréschen Beweismethode erblicken — *aus dem bei weitem leichter zu beweisenden Verhalten der Integrale $J^{(n)}$ bei wachsendem Index Rückschlüsse auch auf das Verhalten der Glieder linker Hand* zu ziehen, aus denen sich die Neumannsche Reihe (Lösung für das Innengebiet) zusammensetzt.

---

[4]) Hilbert hat in seinen Vorlesungen Ende der neunziger Jahre den Beweis der Relation (N) von der Voraussetzung einer konvexen Begrenzungsfläche befreit (vgl. Math. Annalen **55**, S. 4), aber immer blieben doch noch die konvexen Flächen die wichtigsten und am leichtesten zu charakterisierenden, deren Konfigurationskonstante kleiner als 1 ist, so daß die Methode im wesentlichen doch auf konvexe Gebiete beschränkt blieb.

[5]) La méthode de Neumann et le probleme de Dirichlet, Acta mathematica **20**, S. 59.

[6]) Zwei kurze Mitteilungen findet man in den Göttinger Nachrichten 1899 (S. 291) und 1902 (S. 242). Wegen der ausführlicheren Darstellung vgl. unten die Angaben auf S. 450.

Auf dieser Grundlage (P) und ähnlichen Abschätzungen im Gebiete der Robinschen Methode aufbauend, gelang es mir dann schon früher (einfacher wie Poincaré) viele der Konvergenzeigenschaften der Neumannschen und vor allem der Robinschen Potentiale herzuleiten. Doch verzichtete ich dabei auf einen nicht unwesentlichen Vorzug des ursprünglich C. Neumannschen Verfahrens, der auch der Poincaréschen Betrachtungsweise eigen war, wenn er hier auch durch Anwendung recht fernliegender Hilfsmittel etwas teuer erkauft erscheint. Ich folgerte nämlich in meinen älteren Arbeiten die Konvergenz *aus dem Verhalten von Faktoren, welche ihrerseits noch abhingen von den vorgeschriebenen Randwerten,* nämlich der Poincaréschen Integrale (z. B. der $J^{(n)}$). Dadurch kommt in die Konvergenzbetrachtungen eine gewisse Unbestimmtheit. Im Gegensatz dazu tritt z. B. bei C. Neumann von vornherein der *rein geometrische Charakter der Konvergenz* deutlich zutage, die Konvergenz wird bei ihm aus dem Verhalten eines rein geometrischen Faktors, eben der Konfigurationskonstanten c_g, erschlossen, während die gerade vorgeschriebenen Randwerte bei seiner obigen Abschätzung (N) nur in einem allen Gliedern gemeinsamen Faktor, nämlich der Schwankung $\varDelta$, auftreten.

Ich werde nun im folgenden den Konvergenzbeweis auch für nichtkonvexe Gebiete von vornherein möglichst *geometrisch* (in diesem Sinne) führen ohne Heranziehung so fernliegender Hilfsmittel, wie sie Poincaré benutzt, lediglich durch konsequenten Ausbau meiner Methode der Polarfunktionen, und ich möchte hier nur den Grundgedanken kurz andeuten: Bei gegebener Fläche σ hängen die Glieder $W_i^{(n)} - W_i^{(n+1)}$ der Neumannschen Reihe (für das Innengebiet) augenscheinlich ab von den vorgeschriebenen Randwerten f_s und sodann noch von der Lage des Aufpunktes i. Die Methode der Polarfunktionen lieferte mir nun eine Abschätzung dieser Glieder durch das *Produkt zweier Faktoren, deren einer* (J) *immer nur von den Randwerten, der andere* ($\mathfrak{l}_{(i)}$) *nur von der Lage von* i *abhing*:

$$\operatorname{abs}\left(W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+1)}\right) \leqq \frac{1}{4\pi}\sqrt{\mathfrak{l}_{(i)}^{(2\lambda)}}\sqrt{J^{(2\varkappa)}}$$

Diese Faktoren $\mathfrak{l}_{(i)}^{(2\lambda)}$ wären danach also etwa zu bezeichnen als rein *geometrische Funktionen*, nämlich Funktionen der Lage eines Punktes, die im übrigen nur von der vorgegebenen Fläche σ abhängen. — Von diesen Funktionen stellte sich nun heraus, daß sie (unter ziemlich weiten Voraussetzungen über die Fläche σ) für $\lambda \geqq 1$ stetig bleiben selbst bei beliebiger Annäherung von i an σ („stetig in $\mathfrak{J} + \sigma$"), und so führte ich denn früher das Maximum von $\mathfrak{l}_{(i)}''$ ein, das dann natürlich eine rein geometrische Konstante ist, und gelangte so (für $\lambda = 1$ und $\varkappa = n - 1$) zu der obigen

Abschätzung (P), wie sie in den Poincaréschen Betrachtungen steckt. — Im folgenden werde ich nun von diesem Verfahren abweichen und gerade den Index $\varkappa$ des zweiten Faktors J festhalten, so daß also die Randwerte f_s nur in einem unveränderten Faktor (wie früher bei (N) in der Schwankung Δ) auftreten, und werde dann zeigen, daß *der rein geometrische Faktor* $\mathfrak{l}_{(i)}^{(2\lambda)}$ mit wachsendem Index λ, und zwar gleichmäßig für alle Lagen von i in $\mathfrak{J}+\sigma$, gegen 0 konvergiert. Indem ich jetzt also *das Schwergewicht auf den rein geometrischen Faktor* lege, beweise ich, daß die Einzelglieder der Neumannschen Reihe zur 0 abnehmen. Aber auch der Beweis für die Konvergenz der ganzen Reihe glückt mittelst rein potentialtheoretischer Überlegungen, wobei schließlich sogar wieder eine äußerst enge Berührung mit dem ursprünglich C. Neumannschen Gedankengang mit der Konfigurationskonstanten hervortritt.

Äußerst einfach folgt aus diesen Betrachtungen dann auch die Konvergenz der *Robinschen Methode zur Lösung der zweiten Randwertaufgabe*, ja diese Methode ist bei meiner Art des Vorgehens kaum von der Neumannschen zu trennen, die Beziehungen zwischen beiden Methoden spielen im folgenden dauernd eine große Rolle. —

Wie aus dem Gesagten hervorgeht, werde ich mich im folgenden vielfach auf meine früheren Untersuchungen zu berufen haben, die in den beiden Jablonowski-Preisschriften „Studien über die Methoden von C. Neumann und G. Robin zur Lösung der beiden Randwertaufgaben der Potentialtheorie“ und „Beiträge zu einzelnen Fragen der höheren Potentialtheorie“ (bei Teubner 1905 und 1912) niedergelegt sind. Ich werde die benutzten Tatsachen aber stets kurz auseinandersetzen, so daß das Folgende auch ohne Kenntnis jener Arbeiten verständlich sein dürfte. Wegen der Beweise im einzelnen — meist handelt es sich um einfache Anwendungen der Greenschen Sätze — werde ich allerdings stets nur auf jene Arbeiten verweisen. (Ich zitiere sie kurz als „Studien“ und „Beiträge“.)

§ 1.

Die sukzessiven Neumannschen und Robinschen Potentiale und ihre einfachsten Konvergenzeigenschaften.

Die Grundlage der Methode des arithmetischen Mittels, gleichsam die Bausteine, aus denen sich die Neumannsche Lösung der *ersten Randwertaufgabe* aufbaut, bildet eine Folge von Doppelbelegungspotentialen, zu denen man in folgender Weise gelangt: Sind auf der geschlossenen Fläche σ, über die wir die Voraussetzung machen, daß sie überall eine bestimmte Tangentialebene und außerdem stetige Biegung und endliche Krümmung

besitzt [Stud. S. 2], die stetigen Randwerte f_s vorgeschrieben, so ist das erste Potential W das einer Doppelbelegung vom Momente $\frac{1}{2\pi} f$, also

$$W_p = \frac{1}{2\pi}\int f_s \frac{\partial \frac{1}{E_p}}{\partial \nu_s} d\sigma = \frac{1}{2\pi}\int f_\sigma (d\sigma)_p \qquad \left(\begin{array}{l} p \text{ beliebiger Raumpunkt} \\ \nu \text{ innere Flächennormale} \end{array}\right),$$

wo $(d\sigma)_p$ das bekannte C. Neumannsche Symbol, die positiv oder negativ genommene scheinbare Größe des Elementes $d\sigma$ von p aus bedeutet. Die Werte dieses Potentiales in Punkten s von σ selbst (bekanntlich das arithmetische Mittel der inneren und äußeren Randwerte W_{is} und W_{as}) bezeichnen wir dann mit f_s' und bilden ein weiteres Potential W', indem wir diese Werte f_s' an die Stelle der früheren f_s treten lassen, und fahren in dieser Weise fort. So kommen wir sukzessive zu dem folgenden Schema „Neumannscher Potentiale $W^{(n)}$ und Flächenfunktionen $f^{(n)}$“:

$$(\mathfrak{N}_0) \quad \left\{ \begin{array}{ll} W_p = \frac{1}{2\pi}\int f_s (d\sigma)_p, & \frac{1}{2}(W_{is} + W_{as}) = \frac{1}{2\pi}\int f_\sigma (d\sigma)_s = f_s', \\ W_p' = \frac{1}{2\pi}\int f_\sigma' (d\sigma)_p, & \frac{1}{2}(W_{is}' + W_{as}') = \frac{1}{2\pi}\int f_\sigma' (d\sigma)_s = f_s'', \\ W_p'' = \frac{1}{2\pi}\int f_\sigma'' (d\sigma)_p, & \frac{1}{2}(W_{is}'' + W_{as}'') = \frac{1}{2\pi}\int f_\sigma'' (d\sigma)_s = f_s''', \\ \text{usw.} & \text{usw.} \end{array} \right.$$

Dann ist die (zunächst nur bei konvexer Fläche σ als gültig nachgewiesene) Lösung der ersten Randwertaufgabe [7]) für das Innen- bzw. Außengebiet dargestellt durch die beiden „Neumannschen Reihen“:

$$(\mathfrak{N}_1) \quad \begin{array}{l} \Phi_i = C + (W_i - W_i') + (W_i'' - W_i''') + \dots, \\ \Psi_a = C - (W_a + W_a') - (W_a'' + W_a''') - \dots, \end{array}$$

wenn C eine Konstante bedeutet, der sich, wie man beweisen kann, die Funktionen $f^{(n)}$ mit wachsendem n mehr und mehr nähern [vgl. die Einleitung].

Analog geht G. Robin bei der Lösung der *zweiten Randwertaufgabe* vor. An Stelle der *Doppel*belegungspotentiale $W^{(n)}$ treten Potentiale *einfacher* Flächenbelegungen, und die Rolle der Randwerte nehmen die normalen Ableitungen ein. Bezeichnen wir die ursprünglich (für die normalen Ableitungen) vorgegebenen Werte mit g_s, so erhalten wir das folgende Schema „Robinscher Potentiale $V^{(n)}$ und Flächenfunktionen $g^{(n)}$“:

---

[7]) Wegen der näheren Definition vgl. Studien S. 18.

$$(\mathfrak{R}_0)\quad \begin{cases} V_p = \frac{1}{2\pi}\int g_\sigma \frac{d\sigma}{E_p}, & \frac{1}{2}\left[\left(\frac{\partial V}{\partial \nu}\right)_{is} + \left(\frac{\partial V}{\partial \nu}\right)_{as}\right] = \frac{1}{2\pi}\int g_\sigma [d\sigma]_s = g_s', \\ V_p' = \frac{1}{2\pi}\int g_\sigma' \frac{d\sigma}{E_p}, & \frac{1}{2}\left[\left(\frac{\partial V'}{\partial \nu}\right)_{is} + \left(\frac{\partial V'}{\partial \nu}\right)_{as}\right] = \frac{1}{2\pi}\int g_\sigma' [d\sigma]_s = g_s'', \\ V_p'' = \frac{1}{2\pi}\int g_\sigma'' \frac{d\sigma}{E_p}, & \frac{1}{2}\left[\left(\frac{\partial V''}{\partial \nu}\right)_{is} + \left(\frac{\partial V''}{\partial \nu}\right)_{as}\right] = \frac{1}{2\pi}\int g_\sigma'' [d\sigma]_s = g_s''', \\ \text{usw.} & \text{usw.} \end{cases}$$

wo das Symbol $[d\sigma]_s$ in der Bedeutung von $\frac{\partial \frac{1}{E_s}}{\partial \nu_s} d\sigma$ steht $\Big($während $(d\sigma)_s = \frac{\partial \frac{1}{E_s}}{\partial \nu_\sigma} d\sigma$ war!$\Big)$. — Die Flächenbelegungen, von denen die Potentiale V, V', V'' usw. herrühren, haben dann sämtlich die gleiche Masse:

$$(\mathfrak{R}_0') \qquad \frac{1}{2\pi}\int g_\sigma^{(n)} d\sigma = \frac{1}{2\pi}\int g_\sigma d\sigma = m \qquad (n = 1, 2, 3, \ldots).$$

Die Lösungen der zweiten Randwertaufgabe sind dann (wie allerdings in voller Allgemeinheit auch erst noch zu beweisen ist) dargestellt durch die beiden „Robinschen Reihen“:

$$(\mathfrak{R}_1)\quad \begin{aligned} Q_i &= -(V_i + V_i') - (V_i'' - V_i''') - \ldots \\ P_a &= -m\frac{\Pi_a}{2} - (V_a - V_a') - (V_a'' - V_a''') - \ldots \end{aligned}$$

(Π = Pot. d. natürl. Belegung, vgl. später).

Dabei sind beim Innengebiet die Werte g_s aus bekannten Gründen der Beschränkung $m = \frac{1}{2\pi}\int g_s d\sigma = 0$ unterworfen anzunehmen. —

Für diese Neumannschen und Robinschen Potentiale $W^{(n)}$ und $V^{(n)}$ hat nun Poincaré in Anlehnung an H. A. Schwarz merkwürdige *Integraleigenschaften* bewiesen: Kombiniert man z. B. irgend zwei Neumannsche Doppelbelegungspotentiale $W^{(\varkappa)}$ und $W^{(\lambda)}$ in einem Integrale von der Form

$$(1) \qquad [\varphi, \psi] = \int\left(\frac{\partial \varphi}{\partial x}\frac{\partial \psi}{\partial x} + \frac{\partial \varphi}{\partial y}\frac{\partial \psi}{\partial y} + \frac{\partial \varphi}{\partial z}\frac{\partial \psi}{\partial z}\right) d\tau,$$

hinerstreckt entweder über das Innen- oder Außengebiet von σ oder auch über das Totalgebiet $\mathfrak{T} = \mathfrak{J} + \mathfrak{A}$, so hängt der Wert eines solchen Integrals nie von den Einzelindizes $\varkappa$ und λ ab, sondern nur von ihrer Summe $\varkappa + \lambda$ [Stud. S. 75]. Wir setzen daher etwa

$$(2\text{n}) \qquad [W^{(\varkappa)}, W^{(\lambda)}] = J^{(\varkappa+\lambda)}, \quad \text{oder auch ausführlicher} \quad = J^{(\varkappa+\lambda)}(f),$$

um anzudeuten, daß diese Integrale $J^{(n)}$ noch von den vorgegebenen Randwerten f_s abhängen. — Dabei sei noch bemerkt, daß die einfachen Symbole $J^{(n)}$ stets die über das Totalgebiet $\mathfrak{T}$ hinerstreckten Integrale

bedeuten sollen; wo ausnahmsweise einmal Integrale nur über das Innen- oder Außengebiet auftreten, soll das durch ein- bzw. zweimalige Überstreichung des J angedeutet werden [Stud. S. 73].

Analog gilt im Gebiete der Robinschen Methode:

(2r) $$[V^{(\varkappa)}, V^{(\lambda)}] = L^{(\varkappa+\lambda)} \quad \text{oder auch} \quad = L^{(\varkappa+\lambda)}(g)$$

mit entsprechenden Festsetzungen.

Alle Größen $J^{(n)}$ und $L^{(n)}$ mit *geradem* Index n sind augenscheinlich (weil auch darstellbar als Integrale von Quadratsummen) *positiv*. Darüber hinaus läßt sich aber auch leicht zeigen, daß sie zwei Reihen *dauernd abnehmender Größen* bilden

(3) $$\begin{gathered}(J \geqq J'' \geqq) J^{\mathrm{IV}} \geqq J^{\mathrm{VI}} \geqq \ldots \geqq 0 \\ L \geqq L'' \geqq L^{\mathrm{IV}} \geqq L^{\mathrm{VI}} \geqq \ldots \geqq 0\end{gathered} \qquad \text{[Stud. S. 78],}$$

woraus folgt, daß bei beliebigen Ausgangsfunktionen f bzw. g

(3′) $$\lim_{\nu \to \infty} J^{(2\nu)} \quad \textit{und} \quad \lim_{\nu \to \infty} L^{(2\nu)}$$

stets existieren.

Diese Schwarz-Poincaréschen Integraleigenschaften (2) habe ich nun früher in folgender Weise verallgemeinert: Wie wir uns oben z. B. aus den Randwerten f_s die Reihe der Potentiale W, W', W'' usw. hergeleitet dachten, genau so seien jetzt aus anderen (ebenfalls stetigen) Randwerten $\mathfrak{f}_s$ eine neue Folge von Potentialen $\mathfrak{W}$, $\mathfrak{W}'$, $\mathfrak{W}''$ usw. abgeleitet. Dann gilt auch noch für Integrale von der Form $[\varphi, \psi]$, deren beide Bestandteile aus den *verschiedenen* Reihen der W, W', W'' usw. und der $\mathfrak{W}$, $\mathfrak{W}'$, $\mathfrak{W}''$ usw. gewonnen sind, die Eigenschaft, nur von der *Summe der Indizes* der beiden Potentiale abzuhängen:

(4) $$[W^{(\varkappa)}, \mathfrak{W}^{(\lambda)}] = K^{(\varkappa+\lambda)}(f, \mathfrak{f}) \quad \text{und analog} \quad [V^{(\varkappa)}, \mathfrak{V}^{(\lambda)}] = M^{(\varkappa+\lambda)}(g, \mathfrak{g})$$

[Stud. S. 73]. Auch gelten für diese allgemeineren Integrale ebenfalls noch gewisse einfache, im Spezialfalle $\mathfrak{f}_s = f_s$ (bzw. $\mathfrak{g}_s = g_s$) auch schon von Poincaré bewiesene Relationen:

(5) $$\overline{K}^{(n)} - \overline{\overline{K}}^{(n)} = \overline{K}^{(n+1)} + \overline{\overline{K}}^{(n+1)} \quad \text{und} \quad \overline{M}^{(n)} - \overline{\overline{M}}^{(n)} = -(\overline{M}^{(n+1)} + \overline{\overline{M}}^{(n+1)})$$

[Stud. S. 74]. — Diese verallgemeinerten Integrale $K^{(n)}$ und $M^{(n)}$ habe ich nun nachträglich wieder spezialisiert, allerdings nach einer ganz anderen Richtung: Ich machte die eine der beiden Randwertfunktionen, $\mathfrak{f}_s$ bei der Neumannschen Methode, $\mathfrak{g}_s$ bei der Robinschen, abhängig von der Lage eines Punktes (Poles), indem ich z. B. bei innerer Lage dieses Poles (i)

$$\mathfrak{f}_s = \varphi_{(i)s} = \frac{1}{E_{(i)s}} \quad \text{bzw.} \quad \mathfrak{g}_s = \gamma_{(i)s} = \frac{\partial \frac{1}{E_{(i)}}}{\partial \nu_s}$$

setzte. Dann werden auch alle Potentiale $\mathfrak{W}^{(n)}(\mathfrak{V}^{(n)})$ und Flächenfunktionen $\mathfrak{f}^{(n)}(\mathfrak{g}^{(n)})$ abhängig von der Lage dieses Poles i und mit ihnen auch die sämtlichen Integrale $K^{(n)}$ und $M^{(n)}$. Wir setzen daher

$$(6) \qquad K^{(n)}(f, \varphi_{(i)}) = \mathsf{K}_{(i)}^{(n)}(f) \quad \text{bzw.} \quad M^{(n)}(g, \gamma_{(i)}) = \mathsf{M}_{(i)}^{(n)}(g),$$

und, was diese Abhängigkeit einmal natürlich von den Randwerten f_s (bzw. g_s) und sodann eben von der Lage von i anlangt, so stellt sich heraus, daß diese Integrale aufs engste zusammenhängen mit den ausgehend von den Werten $f_s(g_s)$ gebildeten Potentialen $W^{(n)}(V^{(n)})$, es ergeben sich nämlich die fundamentalen Formeln:

$$(7\,i) \qquad \mathsf{K}_{(i)}^{(n)}(f) = 4\pi(W_i^{(n)} - W_i^{(n+1)}), \qquad \mathsf{M}_{(i)}^{(n)}(g) = 4\pi(V_i^{(n)} + V_i^{(n+1)})$$

[Stud. S. 89], und bei äußerer Lage des Poles (a) ganz entsprechend:

$$(7\,a) \qquad \mathsf{K}_{(a)}^{(n)}(f) = -4\pi(W_a^{(n)} + W_a^{(n+1)}), \qquad \mathsf{M}_{(a)}^{(n)}(g) = -4\pi(V_a^{(n)} - V_a^{(n+1)}).$$

Für die noch von der Lage eines Poles o abhängigen, den Ausgangswerten $\varphi_{(o)s} = \frac{1}{E_{os}}$ bzw. $\gamma_{(o)s} = \frac{\partial \frac{1}{E_o}}{\partial \nu_s}$ zugehörigen Potentiale (die wir oben an die Stelle der $\mathfrak{W}^{(n)}$ bzw. $\mathfrak{V}^{(n)}$ treten ließen) wollen wir noch besondere Bezeichnungen einführen, ebenso für die zugehörigen Flächenfunktionen. Es gehören

$$(8) \qquad \begin{aligned} &\text{zu } \varphi_{(o)s} = \frac{1}{E_{os}}: \ \mathsf{W}_{(o)}, \mathsf{W}'_{(o)}, \mathsf{W}''_{(o)} \text{ usw. und } \varphi'_{(o)}, \varphi''_{(o)}, \varphi'''_{(o)} \text{ usw.,} \\ &\text{zu } \gamma_{(o)s} = \frac{\partial \frac{1}{E_o}}{\partial \nu_s}: \ \mathsf{V}_{(o)}, \mathsf{V}'_{(o)}, \mathsf{V}''_{(o)} \text{ usw. und } \gamma'_{(o)}, \gamma''_{(o)}, \gamma'''_{(o)} \text{ usw.} \end{aligned}$$

Dieses sind die *„Polarfunktionen“* (Polarpotentiale und polaren Flächenfunktionen) der Neumannschen bzw. Robinschen Methode, und die ihnen zugehörigen Schwarz-Poincaréschen Integrale $J^{(n)}(\varphi_{(o)})$ bzw. $L^{(n)}(\gamma_{(o)})$ mögen „Polarintegrale“ heißen und mit $\mathsf{I}_{(o)}^{(n)}$ bzw. $\Lambda_{(o)}^{(n)}$ bezeichnet werden, so daß also nach (2):

$$(8') \qquad \mathsf{I}_{(o)}^{(\varkappa+\lambda)} = [\mathsf{W}_{(o)}^{(\varkappa)}, \mathsf{W}_{(o)}^{(\lambda)}] \quad \text{bzw.} \quad \Lambda_{(o)}^{(\varkappa+\lambda)} = [\mathsf{V}_{(o)}^{(\varkappa)}, \mathsf{V}_{(o)}^{(\lambda)}]$$

ist. Diese Integrale hängen dann von keinen irgendwie willkürlich vorgeschriebenen Randwerten ab, sondern sind, wenn eine Fläche σ und die Lage eines Poles o gegeben sind, durch fest vorgeschriebene Regeln definiert. Als nur von geometrischen Verhältnissen abhängige Größen kann man sie etwa als *rein geometrische Funktionen* (der Lage des Poles o) bezeichnen.

Ausführlicher geschrieben lautet z. B. die erste der Gleichungen (7i):

$$(7') \qquad 4\pi(W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+1)}) = [\mathsf{W}_{(i)}^{(\lambda)}, W^{(\varkappa)}]_{\mathfrak{x}},$$

und hieraus folgt nach der Schwarzschen Ungleichung $[\varphi,\psi]^2 \leqq [\varphi,\varphi]\cdot[\psi,\psi]$ unter Benutzung der eingeführten Bezeichnungen:

$$4\pi\cdot \mathrm{abs}\,(W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+1)}) \leqq \sqrt{[\mathsf{W}_{(i)}^{(\lambda)}, \mathsf{W}_{(i)}^{(\lambda)}]\cdot[W^{(\varkappa)}, W^{(\varkappa)}]} = \sqrt{\mathsf{I}_{(i)}^{(2\lambda)}}\cdot\sqrt{J^{(2\varkappa)}(f)},$$

und ganz entsprechende Schlüsse gestatten auch die übrigen Formeln (7), so daß wir erhalten:

$$(9)\quad \begin{cases} |W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+1)}| \leqq \dfrac{1}{4\pi}\sqrt{\mathsf{I}_{(i)}^{(2\lambda)}}\cdot\sqrt{J^{(2\varkappa)}(f)}, \\ |W_a^{(\varkappa+\lambda)} + W_a^{(\varkappa+\lambda+1)}| \leqq \dfrac{1}{4\pi}\sqrt{\mathsf{I}_{(a)}^{(2\lambda)}}\cdot\sqrt{J^{(2\varkappa)}(f)}, \\ |V_i^{(\varkappa+\lambda)} + V_i^{(\varkappa+\lambda+1)}| \leqq \dfrac{1}{4\pi}\sqrt{\Lambda_{(i)}^{(2\lambda)}}\cdot\sqrt{L^{(2\varkappa)}(g)}, \\ |V_a^{(\varkappa+\lambda)} - V_a^{(\varkappa+\lambda+1)}| \leqq \dfrac{1}{4\pi}\sqrt{\Lambda_{(a)}^{(2\lambda)}}\cdot\sqrt{L^{(2\varkappa)}(g)}. \end{cases}$$

Damit sind die schon in der Einleitung erwähnten Abschätzungen gewonnen, ***Abschätzungen der Glieder der Neumannschen und Robinschen Reihen*** ($\mathfrak{R}_1$) ***und*** ($\mathfrak{R}_1$) ***durch Produkte, deren einer Faktor bei gegebener Fläche*** σ ***immer lediglich von dem Aufpunkt*** i ***bzw.*** a ***abhängt, der andere allein von den vorgeschriebenen Randwerten*** f_s (***bzw.*** g_s).

In meinen Studien (Abschn. V) habe ich nun bewiesen, daß die geometrischen Funktionen $\mathsf{I}_{(o)}^{(2\lambda)}$ und $\Lambda_{(o)}^{(2\lambda)}$ unter den über die Fläche σ gemachten Voraussetzungen für $\lambda = 1$ (und erst recht dann auch für größere Werte von λ) stetig bleiben auch bei Annäherung des Poles o an die Fläche σ, so daß sich für sie endliche obere Grenzen angeben lassen:

$$(10)\quad \mathsf{I}''_{(o)} \leqq (4\pi\,\mathsf{A}_g)^2 \quad\text{und}\quad \Lambda''_{(o)} \leqq (4\pi\,\mathsf{B}_g)^2 \quad (o \text{ beliebig in } \mathfrak{J}+\sigma \text{ oder } \mathfrak{A}+\sigma).$$

Die Konstanten A_g und B_g hängen dann nur von den geometrischen Verhältnissen der Fläche σ ab. Ihre Einführung in (9) liefert (für $\varkappa = n-1$) die Abschätzungen:

$$\left.\begin{matrix} |W_i^{(n)} - W_i^{(n+1)}| \\ |W_a^{(n)} + W_a^{(n+1)}| \end{matrix}\right\} \leqq \mathsf{A}_g\sqrt{J^{(2n-2)}}, \qquad \left.\begin{matrix} |V_i^{(n)} + V_i^{(n+1)}| \\ |V_a^{(n)} - V_a^{(n+1)}| \end{matrix}\right\} \leqq \mathsf{B}_g\sqrt{L^{(2n-2)}}$$

und damit, wie schon in der Einleitung erwähnt wurde, die Grundlage meiner älteren Arbeiten und den Anschluß an die Poincaréschen Methoden.

Jetzt aber wollen wir einen anderen Weg einschlagen, wobei es allerdings zweckmäßig ist, an Stelle der über das Totalgebiet $\mathfrak{T} = \mathfrak{J} + \mathfrak{A}$ hinerstreckten Integrale (J, K bzw. L, M) die über die Einzelgebiete $\mathfrak{J}$ bzw. $\mathfrak{A}$ erstreckten, durch ein- bzw. zweimalige Überstreichung gekennzeichneten Integrale in den Vordergrund zu stellen: Es ist allgemein

$$\overline{K}^{(n)} + \overline{\overline{K}}^{(n)} = K^{(n)}$$

und andererseits folgt aus (5):

$$\overline{K}^{(n)} - \overline{\overline{K}}^{(n)} = \overline{K}^{(n+1)} + \overline{\overline{K}}^{(n+1)} = K^{(n+1)}$$

und daher

$$\overline{K}^{(n)} = \frac{1}{2}(K^{(n)} + K^{(n+1)}) \quad \text{und} \quad \overline{\overline{K}}^{(n)} = \frac{1}{2}(K^{(n)} - K^{(n+1)}).$$

Wenden wir diese Formeln nun auf den Fall an, daß die einen in den Integralen K steckenden Potentiale $\mathfrak{W}$ [vgl. (4)] die Polarpotentiale $\mathsf{W}_{(i)}$ bzw. $\mathsf{W}_{(a)}$ sind, so folgt nach (7):

$$\overline{\mathsf{K}}_{(i)}^{(n)}(f) = 2\pi(W_i^{(n)} - W_i^{(n+1)}) + 2\pi(W_i^{(n+1)} - W_i^{(n+2)}) = 2\pi(W_i^{(n)} - W_i^{(n+2)})$$

und analog:

$$\overline{\overline{\mathsf{K}}}_{(a)}^{(n)}(f) = -2\pi(W_a^{(n)} - W_a^{(n+2)})$$

oder ausführlicher:

$$(11) \qquad \begin{aligned} 2\pi(W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+2)}) &= [\mathsf{W}_{(i)}^{(\lambda)}, W^{(\varkappa)}]_{\mathfrak{J}} \\ 2\pi(W_a^{(\varkappa+\lambda)} - W_a^{(\varkappa+\lambda+2)}) &= -[\mathsf{W}_{(a)}^{(\lambda)}, W^{(\varkappa)}]_{\mathfrak{A}} \end{aligned}$$

und daraus wieder nach der Schwarzschen Ungleichung:

$$(12) \qquad \begin{aligned} |W_i^{(\varkappa+\lambda)} - W_i^{(\varkappa+\lambda+2)}| &\leqq \frac{1}{2\pi}\sqrt{\overline{\mathsf{I}}_{(i)}^{(2\lambda)}} \cdot \sqrt{\overline{J}^{(2\varkappa)}(f)} \\ |W_a^{(\varkappa+\lambda)} - W_a^{(\varkappa+\lambda+2)}| &\leqq \frac{1}{2\pi}\sqrt{\overline{\mathsf{I}}_{(a)}^{(2\lambda)}} \cdot \sqrt{\overline{\overline{J}}^{(2\varkappa)}(f)}. \end{aligned}$$

Damit haben wir Formeln von derselben Art wie oben in (9) erhalten, nur daß jetzt Integrale nur über das Innengebiet $\mathfrak{J}$ oder über das Außengebiet $\mathfrak{A}$ auftreten. — Jetzt halten wir aber im Gegensatz zu früher $\varkappa$ fest, nämlich gleich 2 (die angenommene Stetigkeit der Randwerte f_s genügt, die Endlichkeit bereits von $\overline{J}^{\text{IV}}$ und $\overline{\overline{J}}^{\text{IV}}$ zu verbürgen, Stud. S. 73). Dann erhalten wir Abschätzungen, in denen die Randwerte f_s nur in einem bei wachsendem Index ($n = \varkappa + 2$) stets gleichbleibenden Faktor enthalten sind, also Abschätzungen im wesentlichen wieder durch rein geometrische Funktionen:

$$(13) \qquad \begin{aligned} |W_i^{(n)} - W_i^{(n+2)}| &\leqq \frac{1}{2\pi}\sqrt{\overline{J}^{\text{IV}}} \cdot \sqrt{\overline{\mathsf{I}}_{(i)}^{(2n-4)}} \\ |W_a^{(n)} - W_a^{(n+2)}| &\leqq \frac{1}{2\pi}\sqrt{\overline{\overline{J}}^{\text{IV}}} \cdot \sqrt{\overline{\mathsf{I}}_{(a)}^{(2n-4)}}. \end{aligned}$$

Um nun weiterzugehen, knüpfen wir wieder an die Formeln (5) an, wenden sie aber jetzt nur auf den von Poincaré allein betrachteten Fall an, daß die beiden in jedem der Integrale K steckenden Potentiale [vgl. (4)] demselben Systeme angehören ($\mathfrak{f}_s = f_s$), also an die Stelle unserer allgemeinen Integrale $K^{(n)}(f, \mathfrak{f})$ die spezielleren Integrale $J^{(n)}(f)$ treten [vgl. (2)] und analog $L^{(n)}(g)$ an Stelle von $M^{(n)}(g, \mathfrak{g})$. — Dann folgt aus jenen Formeln (5) leicht:

$$\begin{aligned} 2J^{(2\nu+1)} &= \quad (J^{(2\nu)} + J^{(2\nu+2)}) - (\overline{J}^{(2\nu)} - \overline{J}^{(2\nu+2)}) \\ 2L^{(2\nu+1)} &= -(L^{(2\nu)} + L^{(2\nu+2)}) + (\overline{L}^{(2\nu)} - \overline{L}^{(2\nu+2)}) \end{aligned} \qquad [\text{Stud. S. 77}].$$

Diese Formeln wenden wir nun speziell auf unsere Polarpotentiale an und machen noch Gebrauch von der früher [Stud. S. 121] bewiesenen Tatsache, daß entsprechende Polarpotentiale $\mathsf{W}_{(i)}^{(n)}$ und $\mathsf{V}_{(i)}^{(n)}$ der Neumannschen und Robinschen Methode bei innerer Lage des Poles im ganzen Innengebiete $\mathfrak{J}$ übereinstimmen, und daß daher auch die über das *Innengebiet* hinerstreckten Polarintegrale $\overline{\mathsf{I}}_{(i)}^{(n)}$ und $\overline{\Lambda}_{(i)}^{(n)}$ gleich sind. Dann folgt sofort:

$$(14\mathrm{i}) \qquad 2(\overline{\mathsf{I}}_{(i)}^{(2\nu)} + \overline{\mathsf{I}}_{(i)}^{(2\nu+2)}) = (\overline{\overline{\mathsf{I}}}_{(i)}^{(2\nu)} - \overline{\overline{\mathsf{I}}}_{(i)}^{(2\nu+2)}) + (\overline{\overline{\Lambda}}_{(i)}^{(2\nu)} - \overline{\overline{\Lambda}}_{(i)}^{(2\nu+2)})$$

und, da die Größen $\overline{\overline{\mathsf{I}}}_{(i)}^{(2\nu)}$ und $\overline{\overline{\Lambda}}_{(i)}^{(2\nu)}$ als positive und mit wachsendem ν stets abnehmende Größen sich bestimmten Grenzwerten nähern [vgl. (3)], so folgt hieraus das wichtige Resultat:

$$(15\mathrm{i}) \qquad \lim_{\nu \to \infty} \overline{\mathsf{I}}_{(i)}^{(2\nu)} = 0,$$

und, da alle diese Einzelfunktionen $\overline{\mathsf{I}}_{(i)}^{(2\nu)}$, wie schon erwähnt, innerhalb $\mathfrak{J} + \sigma$ stetig sind, und die Abnahme gegen 0 nach (3) für jedes i *monoton* erfolgt, so ergibt sich weiter nach einem bekannten Dinischen Satze, daß diese Konvergenz auch für alle Lagen von i im Innengebiet eine *gleichmäßige* ist, ja sogar, daß die Maxima von $\overline{\mathsf{I}}_{(i)}^{(2\nu)}$ im Gebiete $\mathfrak{J} + \sigma$ eine monoton zur 0 abnehmende Reihe von Größen bilden — und ganz entsprechende Überlegungen führen über die Formel

$$(14\mathrm{a}) \qquad 2(\overline{\mathsf{I}}_{(a)}^{(2\nu)} + \overline{\mathsf{I}}_{(a)}^{(2\nu+2)}) = (\overline{\mathsf{I}}_{(a)}^{(2\nu)} - \overline{\mathsf{I}}_{(a)}^{(2\nu+2)}) + (\overline{\Lambda}_{(a)}^{(2\nu)} - \overline{\Lambda}_{(a)}^{(2\nu+2)})$$

zu dem Resultate, daß auch

$$(15\mathrm{a}) \qquad \lim_{\nu \to \infty} \overline{\mathsf{I}}_{(a)}^{(2\nu)} = 0$$

ist und zwar *gleichmäßig* im ganzen Gebiete $\mathfrak{A} + \sigma$.[8])

---

[8]) Man braucht den Beweis zunächst nur für ein *abgeschlossenes* Gebiet zu führen, etwa wie es begrenzt wird von der Fläche σ und einer σ umschließenden Hilfsfläche σ_0, denn für die außerhalb σ_0 gelegenen Punkte a ergibt sich die Richtigkeit unserer Behauptung aus der Tatsache, daß das Maximum von $\overline{\mathsf{I}}_{(a)}^{(2\nu)}$ im Gebiete $\mathfrak{A}_0 + \sigma_0$ speziell auf σ_0 anzutreffen ist. — Das folgt leicht in folgender Weise (mittelst einer Darstellung von $\overline{\mathsf{I}}_{(a)}^{(2\nu)}$ als Wert eines Potentials nach (11) und Anwendung bekannter Sätze über Extrema von Potentialen):

$$\overline{\mathsf{I}}_{(a)}^{(2\nu)} = 2\pi(\mathsf{W}_{(a)a}^{(2\nu+2)} - \mathsf{W}_{(a)a}^{(2\nu)}) \leq 2\pi(\mathsf{W}_{(a)s_0}^{(2\nu+2)} - \mathsf{W}_{(a)s_0}^{(2\nu)}) \leq \sqrt{\overline{\mathsf{I}}_{(a)}^{(2\nu)}\,\overline{\mathsf{I}}_{(s_0)}^{(2\nu)}} \qquad [\text{vgl. (12)}],$$

also

$$\overline{\mathsf{I}}_{(a)}^{(2\nu)} \leq \overline{\mathsf{I}}_{(s_0)}^{(2\nu)},$$

wenn a ein *beliebiger* Punkt *außerhalb* σ_0 und s_0 ein *geeignet* gewählter Punkt *auf* σ_0 ist. — Q. e. d.

Nach diesen Feststellungen (15i) und (15a) liefert uns nun (13) sofort die Resultate

$$(16)\qquad \lim_{n\to\infty}(W_i^{(n)} - W_i^{(n+2)}) = 0 \quad \text{und} \quad \lim_{n\to\infty}(W_a^{(n)} - W_a^{(n+2)}) = 0$$

und zwar wieder *gleichmäßig* in den Gebieten $\mathfrak{J} + \sigma$ bzw. $\mathfrak{A} + \sigma$.

Da aber nach dem allgemeinen auf S. 451 gegebenen Schema $(\mathfrak{N}_0)$ $W^{(n)} - W^{(n+2)}$ das Potential einer Doppelbelegung vom Momente $\frac{1}{2\pi}(f^{(n)} - f^{(n+2)})$ ist, und das Moment stets die halbe Differenz der innern und äußern Randwerte des Potentials darstellt, so folgt aus (16), daß auch

$$\lim_{n\to\infty}(f_s^{(n)} - f_s^{(n+2)}) = 0$$

ist, und zwar *gleichmäßig* für alle Lagen von s auf σ.

Nun sind aber die Werte $f_s^{(n)} - f_s^{(n+2)}$ zugleich die inneren Randwerte des Potentials $W_i^{(n)} - W_i^{(n+1)}$ und die (negativen) äußeren von $W_a^{(n)} + W_a^{(n+1)}$ [vgl. Schema $(\mathfrak{N}_0)$], und da unter diesen Randwerten bekanntlich die größten und kleinsten *aller* Werte anzutreffen sind, so gilt also bei beliebiger Lage von i bzw. a

$$(17)\qquad \lim_{n\to\infty}(W_i^{(n)} - W_i^{(n+1)}) = 0 \quad \text{und} \quad \lim_{n\to\infty}(W_a^{(n)} + W_a^{(n+1)}) = 0,$$

und zwar wieder *gleichmäßig in den Gebieten* $\mathfrak{J} + \sigma$ bzw. $\mathfrak{A} + \sigma$.

Damit ist zunächst der Nachweis erbracht, daß wenigstens *die Glieder der Neumannschen Reihen* $(\mathfrak{N}_1)$ *zur* 0 *abnehmen* — und zwar weit einfacher und direkter als ich diesen Beweis noch in meinen Studien [S. 113] geführt habe, und vor allem erscheint erst jetzt die Konvergenz lediglich bedingt durch rein geometrische Größen und nicht mehr, wie früher, abhängig von den zufällig vorgeschriebenen Randwerten.

Wir wollen dieses bisher rein qualitative Resultat (17) nun zunächst noch benutzen, um die sich ergebenden *Abschätzungen* möglichst einfach zu gestalten: Wir wenden es auf die Polarpotentiale an und lassen Pol und Aufpunkt zusammenfallen. Dann folgt, daß bei festgehaltenem i oder a sicher

$$\lim_{n\to\infty}(\mathsf{W}_{(i)i}^{(n)} - \mathsf{W}_{(i)i}^{(n+1)}) = 0 \quad \text{und} \quad \lim_{n\to\infty}(\mathsf{W}_{(a)a}^{(n)} + \mathsf{W}_{(a)a}^{(n+1)}) = 0$$

ist. Nun ergibt sich aber durch Spezialisierung aus (7) [vgl. auch (7′) und (8′)]:

$$4\pi(\mathsf{W}_{(i)i}^{(2\nu)} - \mathsf{W}_{(i)i}^{(2\nu+1)}) = \mathsf{I}_{(i)}^{(2\nu)} \quad \text{und} \quad -4\pi(\mathsf{W}_{(a)a}^{(2\nu)} + \mathsf{W}_{(a)a}^{(2\nu+1)}) = \mathsf{I}_{(a)}^{(2\nu)}$$

[Stud. S. 125], d. h. gleich den *über den Totalraum* $\mathfrak{T} = \mathfrak{J} + \mathfrak{A}$ *hinerstreckten Polarintegralen*, also auch diese nähern sich mit wachsendem

Index der 0, und zwar, da die Abnahme monoton erfolgt [vgl. (3)], auch *gleichmäßig in allen Punkten von* $\mathfrak{J}+\sigma$ *bzw.* $\mathfrak{A}+\sigma$:

$$\lim_{\nu\to\infty} \mathsf{I}_{(i)}^{(2\nu)} = 0 \quad \text{und} \quad \lim_{\nu\to\infty} \mathsf{I}_{(a)}^{(2\nu)} = 0.$$

Setzen wir also immer noch den größten Wert von $\mathsf{I}_{(o)}^{(2\nu-4)}$ bei beliebiger Lage des Poles in $\mathfrak{J}+\sigma$ oder $\mathfrak{A}+\sigma$, das

$$\text{(18)} \qquad \text{Max.}\, \mathsf{I}_{(o)}^{(2\nu-4)} = (4\pi\varepsilon_\nu)^2 \qquad (\nu \geqq 3),$$

so bilden *die* ε_ν *eine Reihe* ***rein geometrischer***, *d. h. allein der Fläche* σ *eigentümlicher, Konstanten, die mit wachsendem* ν *monoton gegen* 0 *abnehmen*

$$\text{(18')} \qquad \varepsilon_3 \geqq \varepsilon_4 \geqq \varepsilon_5 \geqq \dots \to 0$$

und ihre Einführung gestattet dann, den Abschätzungen (9) (für $\varkappa = 2$, $\varkappa+\lambda = n$) die folgende Form zu geben:

$$\text{(17')} \qquad \left.\begin{array}{l} \text{abs}\,(W_i^{(n)} - W_i^{(n+1)}) \\ \text{abs}\,(W_a^{(n)} + W_a^{(n+1)}) \end{array}\right\} \leqq \sqrt{J^{\mathrm{IV}}} \cdot \varepsilon_n,$$

aus der jetzt besonders deutlich der *rein geometrische Charakter der Konvergenz* dieser Glieder gegen 0 ersichtlich ist.

§ 2.

Fortsetzung. Insbesondere über die Konvergenzeigenschaften der Robinschen Potentiale.

Zu erheblich weitergehenden Resultaten führte mich die Methode der Polarfunktionen im Gebiete der Robinschen Methode zur Lösung der zweiten Randwertaufgabe [vgl. Stud. S. 114]. Nun gilt es aber noch, unter Beibehaltung des Grundgedankens, die früher gegebenen Beweise so umzugestalten, daß auch hier der geometrische Charakter der Konvergenz deutlich hervortritt. Daher wollen wir uns zunächst nur mit den *Polarpotentialen* $\mathsf{V}_{(o)}^{(n)}$ der Robinschen Methode [vgl. oben (8)] als rein geometrischen Funktionen beschäftigen.

Wir wenden die im Gebiete der Robinschen Potentiale gefundenen Abschätzungen (9) zunächst auf den Fall an, daß $\lambda = 1$ ist und die vorgeschriebenen Werte

$$g_s = \gamma_{(i)s} - \gamma_{(i)s}^{(2q)} \qquad (q \text{ beliebig} = 1, 2, 3, \dots)$$

sind, wo i ein an sich beliebiger, aber zunächst immer *fest* gedachter innerer Punkt sei. Ersetzt man dann noch $\Lambda_{(i)}''$ und $\Lambda_{(a)}''$ durch den zu großen Wert $(4\pi\mathsf{B}_g)^2$ [vgl. (10)], entsprechend dem in meinen Studien

eingeschlagenen Wege, so erhält man unabhängig von der Lage von i oder a:

$$\left.\begin{array}{l}\mathrm{abs}\left(\left(\mathsf{V}_{(\mathrm{i})i}^{(n)}-\mathsf{V}_{(\mathrm{i})i}^{(n+2q)}\right)+\left(\mathsf{V}_{(\mathrm{i})i}^{(n+1)}-\mathsf{V}_{(\mathrm{i})i}^{(n+2q+1)}\right)\right)\\ \mathrm{abs}\left(\left(\mathsf{V}_{(\mathrm{i})a}^{(n)}-\mathsf{V}_{(\mathrm{i})a}^{(n+2q)}\right)-\left(\mathsf{V}_{(\mathrm{i})a}^{(n+1)}-\mathsf{V}_{(\mathrm{i})a}^{(n+2q+1)}\right)\right)\end{array}\right\}\leqq \mathsf{B}_g\sqrt{L^{(2n-2)}\left(\gamma_{(\mathrm{i})}-\gamma_{(\mathrm{i})}^{(2q)}\right)}. \tag{19}$$

Da aber nach dem Bildungsgesetz (2r) der Größen $L^{(n)}$ ganz *allgemein*

$$\begin{aligned}L^{(2\nu)}(g-g^{(2q)}) &= L^{(2\nu)}-2L^{(2\nu+2q)}+L^{(2\nu+4q)}\\ &= (L^{(2\nu)}-L^{(2\nu+2q)})-(L^{(2\nu+2q)}-L^{(2\nu+4q)}),\end{aligned}$$

also nach (3) und (3′)

$$< L^{(2\nu)}-L^{(2\nu+2q)} < L^{(2\nu)}-\lim_{\nu\to\infty}L^{(2\nu)}$$

ist, und daher ***gleichmäßig für alle positiven*** q:

$$\lim_{\nu\to 0}L^{(2\nu)}(g-g^{(2q)})=0,$$

so folgt aus den Abschätzungen (19), daß die Ausdrücke links gleichmäßig für alle Punkte i von $\mathfrak{J}+\sigma$ bzw. a von $\mathfrak{A}+\sigma$ zur 0 konvergieren. In Punkten s auf σ gehen demnach *beide* Ausdrücke gegen 0, und daraus folgt durch Kombination, daß auch

$$\lim_{n\to\infty}\left(\mathsf{V}_{(\mathrm{i})s}^{(n)}-\mathsf{V}_{(\mathrm{i})s}^{(n+2q)}\right)=0 \quad \text{gleichmäßig für alle } s \text{ auf } \sigma \text{ und beliebige positive } q.$$

Da aber unter diesen *Rand*werten der Potentiale $\mathsf{V}_{(\mathrm{i})}^{(n)}-\mathsf{V}_{(\mathrm{i})}^{(n+2q)}$ die absolut größten *aller* Werte vorkommen, so folgt daraus weiter, daß ***sich die Potentiale*** $\mathsf{V}_{(\mathrm{i})p}^{(n)}$ ***mit geradem und ebenso die mit ungeradem Index gleichmäßig im ganzen Raum je einer Konvergenzfunktion nähern****). Aus diesem Sachverhalt ergibt sich dann aber, wie ich in den Studien S. 100 bis 105 gezeigt habe, daß diese beiden Funktionen zunächst ***übereinstimmen,***

---

*) Die *Tatsache* der Konvergenz an sich folgt leicht auch so: Wenn wir beachten, daß bei innerer Lage des Poles die Polarpotentiale der Neumannschen und Robinschen Methode im ganzen Innengebiet identisch sind $\left(\mathsf{W}_{(\mathrm{i})i}^{(n)}=\mathsf{V}_{(\mathrm{i})i}^{(n)}\right.$, Stud. S. 121$\left.\right)$, so ist nach (11) und (12) z. B. für gerades $\varkappa+\lambda=n=2\nu$

$$\begin{aligned}\left|\mathsf{V}_{(\mathrm{i})i}^{(n)}-\mathsf{V}_{(\mathrm{i})i}^{(n+2)}\right| &= \left|\mathsf{W}_{(\mathrm{i})i}^{(n)}-\mathsf{W}_{(\mathrm{i})i}^{(n+2)}\right|\\ &= \frac{1}{2\pi}\left[\mathsf{W}_{(\mathrm{i})}^{(\nu)},\mathsf{W}_{(i)}^{(\nu)}\right]_{\mathfrak{J}} \leqq \frac{1}{2\pi}\sqrt{\bar{\mathrm{I}}_{(\mathrm{i})}^{(2\nu)}\,\bar{\mathrm{I}}_{(i)}^{(2\nu)}} \leqq \frac{1}{4\pi}\left(\bar{\mathrm{I}}_{(\mathrm{i})}^{(2\nu)}+\bar{\mathrm{I}}_{(i)}^{(2\nu)}\right),\end{aligned}$$

und da nach (14i) auch $\sum\bar{\mathrm{I}}_{(i)}^{(2\nu)}$ konvergiert, so folgt die Behauptung durch Anwendung dieser letzteren Formel auf die Fälle n, $n+2$, $n+4$ usw. und Addition — doch ist es mir so nicht gelungen, auch die *Gleichmäßigkeit* der Konvergenz zu beweisen: der Dinische Satz versagt hier, weil die Stetigkeit der Konvergenzwerte von $\sum\bar{\mathrm{I}}_{(i)}^{(2\nu)}$ (d. i. nach (14i) im wesentlichen von $\lim_{\nu\to\infty}\bar{\bar{\Lambda}}_{(i)}^{(2\nu)}$), wenigstens zunächst, noch nicht nachgewiesen ist.

dann aber auch, daß sie im Innengebiete $\mathfrak{J}$ *konstant* sein müssen[10]), doch könnte der Wert dieser Konstanten nach dem Bisherigen noch von der Wahl des Poles $\mathfrak{i}$ abhängen. Wir setzen daher zunächst

$$\lim_{n\to\infty} \mathsf{V}_{(\mathfrak{i})\,i}^{(n)} = C_{\mathfrak{i}}\,.$$

Nach dem allgemeinen Reziprozitätsgesetze der Polarpotentiale:

$$\mathsf{V}_{(\mathfrak{i}_1)\,\mathfrak{i}_2}^{(n)} = \mathsf{V}_{(\mathfrak{i}_2)\,\mathfrak{i}_1}^{(n)} \qquad \text{[Stud. S. 123]}$$

folgt aber, daß $C_{(\mathfrak{i}_1)} = C_{(\mathfrak{i}_2)}$ ist, also dieser konstante Wert von der spezielleren Lage des (inneren) Poles unabhängig und mithin lediglich von den geometrischen Verhältnissen der Fläche σ abhängig sein muß. Deshalb setzen wir:

(20i) $$\lim_{n\to\infty} \mathsf{V}_{(\mathfrak{i})\,i}^{(n)} = 2\,\Gamma_g \quad \text{gleichmäßig für alle } i \text{ in } \mathfrak{J}+\sigma \textit{ bei festem } \mathfrak{i}.$$

Für das Außengebiet kann man entsprechende Betrachtungen anstellen, ja hier ist man sogar in der Lage, den Wert der zugehörigen Konvergenzkonstanten sogleich zu 0 anzugeben (weil nämlich schon alle *einzelnen* $\mathsf{V}_{(\mathfrak{a})}^{(n)}$ als Potentiale von Belegungen mit verschwindender Gesamtmasse den Wert 0 annehmen müssen). Somit folgt

(20a) $$\lim_{n\to\infty} \mathsf{V}_{(\mathfrak{a})\,a}^{(n)} = 0 \quad \text{gleichmäßig für alle } a \text{ in } \mathfrak{A}+\sigma \textit{ bei festem } \mathfrak{a}.$$

Nachträglich läßt sich nun auch zeigen, daß diese Konvergenzen (20i) und (20a) *auch in* $\mathfrak{i}$ *und* $\mathfrak{a}$ (!) *gleichmäßig* sind: Zunächst folgt aus (7) und (8′) durch Spezialisiernng ($g_s = \gamma_{(\mathfrak{i})s}$ bzw. $g_s = \gamma_{(\mathfrak{a})s}$)

$$\Lambda_{(\mathfrak{i})}^{(n)} = 4\pi\left(\mathsf{V}_{(\mathfrak{i})\,i}^{(n)} + \mathsf{V}_{(\mathfrak{i})\,i}^{(n+1)}\right) \quad \text{und} \quad \Lambda_{(\mathfrak{a})}^{(n)} = -4\pi\left(\mathsf{V}_{(\mathfrak{a})\,a}^{(n)} - \mathsf{V}_{(\mathfrak{a})\,a}^{(n+1)}\right),$$

und daher weiter nach (20) für gerades $n = 2\nu$, zunächst bei festem $\mathfrak{i}$ bzw. $\mathfrak{a}$:

(21) $$\lim_{\nu\to\infty} \Lambda_{(\mathfrak{i})}^{(2\nu)} = 16\,\pi\,\Gamma_g \quad \text{und} \quad \lim_{\nu\to\infty} \Lambda_{(\mathfrak{a})}^{(2\nu)} = 0.$$

Da sich aber diese stetigen Funktionen $\Lambda_{(\mathfrak{i})}^{(2\nu)}$ und $\Lambda_{(\mathfrak{a})}^{(2\nu)}$ *monoton* den angegebenen stetigen Grenzfunktionen (Konstanten) nähern [vgl. (3)], so ist auch diese Konvergenz nach Dini wieder eine *gleichmäßige* für alle

---

[10]) Das erstere, daß nämlich jene beiden Konvergenzfunktionen *gleich* sind, folgt, zunächst wenigstens für das Innengebiet (und daraus dann leicht auch für den ganzen Raum), auch aus dem Hauptresultat von § 1. Denn nach (17) ist

$$\lim_{\nu\to\infty}\left(\mathsf{V}_{(\mathfrak{i})\,i}^{(2\nu)} - \mathsf{V}_{(\mathfrak{i})\,i}^{(2\nu+1)}\right) = \lim_{\nu\to\infty}\left(\mathsf{W}_{(\mathfrak{i})\,i}^{(2\nu)} - \mathsf{W}_{(\mathfrak{i})\,i}^{(2\nu+1)}\right) = 0, \quad \text{also:} \quad \lim_{\nu\to\infty} \mathsf{V}_{(\mathfrak{i})\,i}^{(2\nu+1)} = \lim_{\nu\to\infty} \mathsf{V}_{(\mathfrak{i})\,i}^{(2\nu)}$$

– doch zum Beweise der *Konstanz* kann ich vorläufig die in meinen Studien benutzten Sätze von der Konstanz des Momentes einer Doppelbelegung [vgl. daselbst S. 50–51] noch nicht entbehren.

Lagen von i bzw. a, und wir können wieder *eine monoton zur* 0 *abnehmende Folge von rein geometrischen Größen* η_2, η_3 usw. einführen, indem wir allgemein setzen:

$$(22)\quad (4\pi\eta_\nu)^2 = \textit{der größeren der Zahlen} \begin{cases} \text{Max.}\,(\mathsf{\Lambda}_{(i)}^{(2\nu-2)} - 16\pi\mathsf{\Gamma}_g) & \text{in } \mathfrak{J}+\sigma, \\ \text{Max.}\,\mathsf{\Lambda}_{(a)}^{(2\nu-2)} & \text{in } \mathfrak{A}+\sigma, \end{cases}$$

so daß dann also

$$(22')\qquad \eta_2 \geqq \eta_3 \geqq \eta_4 \geqq \ldots \rightarrow 0.$$

Nun enthalten aber die an (19) anknüpfenden Betrachtungen (bzw. die entsprechenden fürs Außengebiet) in Verbindung mit (21) die Abschätzungen:

$$\text{abs}\,(\mathsf{V}_{(\mathfrak{i})i}^{(n)} - 2\,\mathsf{\Gamma}_g) \leqq \mathsf{B}_g \sqrt{\mathsf{\Lambda}_{(\mathfrak{i})}^{(2n-2)} - 16\pi\mathsf{\Gamma}_g} \qquad \text{und} \qquad \text{abs}\,\mathsf{V}_{(\mathfrak{a})a}^{(n)} \leqq \mathsf{B}_g \sqrt{\mathsf{\Lambda}_{(\mathfrak{a})}^{(2n-2)}},$$

und daher ergeben sich die Formeln

$$(23)\qquad \left.\begin{array}{l} \text{abs}\,(\mathsf{V}_{(\mathfrak{i})i}^{(n)} - 2\,\mathsf{\Gamma}_g) \\ \text{abs}\,\mathsf{V}_{(\mathfrak{a})a}^{(n)} \end{array}\right\} \leqq 4\pi\,\mathsf{B}_g\cdot\eta_n,$$

aus denen jetzt *die Gleichmäßigkeit der in* (20i) *und* (20a) *angegebenen Konvergenzen in i und* $\mathfrak{i}$ *bzw. in a und* $\mathfrak{a}$ unmittelbar ersichtlich ist:

$$(20')\qquad \lim_{n\to\infty} \begin{cases} \mathsf{V}_{(\mathfrak{i})i}^{(n)} = 2\,\mathsf{\Gamma}_g & \textit{gleichmäßig} \text{ für alle } i \textit{ und } \mathfrak{i} \text{ von } \mathfrak{J}+\sigma, \\ \mathsf{V}_{(\mathfrak{a})a}^{(n)} = 0 & \quad\text{"}\qquad\text{"}\quad\text{"}\ a\ \text{"}\ \mathfrak{a}\ \text{"}\ \mathfrak{A}+\sigma. \end{cases}$$

Bevor wir dieses *Hauptresultat über die Polarpotentiale* auf die allgemeinen Potentiale der Robinschen Methode übertragen, wollen wir jetzt noch eine für das Folgende nützliche Anwendung machen: Wir bilden die Schwarz-Poincaréschen Integrale $L^{(n)}$ [vgl. (2r)] für den Fall, daß die vorgegebenen Randwerte

$$g_s = \gamma_{(o)s} - \gamma_{(\mathfrak{o})s}^{(q)} \qquad (q = 0, 1, 2, 3, \ldots)$$

sind, wo o und $\mathfrak{o}$ zwei beliebige gleichartige, d. h. beides innere oder beides äußere Punkte sind. Dann ergibt sich, da allgemein nach (2r) und (4)

$$\begin{aligned} L^{(2\nu)}(\gamma_{(o)} - \gamma_{(\mathfrak{o})}^{(q)}) &= L^{(2\nu)}(\gamma_{(o)}) - 2\,M^{(2\nu+q)}(\gamma_{(o)}, \gamma_{(\mathfrak{o})}) + L^{(2\nu+2q)}(\gamma_{(\mathfrak{o})}) \\ &= \mathsf{M}_{(o)}^{(2\nu)}(\gamma_{(o)}) - 2\,\mathsf{M}_{(o)}^{(2\nu+q)}(\gamma_{(\mathfrak{o})}) + \mathsf{M}_{(\mathfrak{o})}^{(2n+2q)}(\gamma_{(\mathfrak{o})}) \end{aligned}$$

ist, mit Rücksicht auf die Formeln (7i) bzw. (7a):

$$\begin{aligned} L^{(2\nu)}(\gamma_{(i)} - \gamma_{(\mathfrak{i})}^{(q)}) = 4\pi\big[&(\mathsf{V}_{(i)i}^{(2\nu)} + \mathsf{V}_{(i)i}^{(2\nu+1)} - 4\,\mathsf{\Gamma}_g) \\ &- 2\,(\mathsf{V}_{(\mathfrak{i})i}^{(2\nu+q)} + \mathsf{V}_{(\mathfrak{i})i}^{(2\nu+q+1)} - 4\,\mathsf{\Gamma}_g) + (\mathsf{V}_{(\mathfrak{i})\mathfrak{i}}^{(2\nu+2q)} + \mathsf{V}_{(\mathfrak{i})\mathfrak{i}}^{(2\nu+2q+1)} - 4\,\mathsf{\Gamma}_g)\big] \end{aligned}$$

bzw.

$$\begin{aligned} L^{(2\nu)}(\gamma_{(a)} - \gamma_{(\mathfrak{a})}^{(q)}) = -\,4\pi\big[&(\mathsf{V}_{(a)a}^{(2\nu)} - \mathsf{V}_{(a)a}^{(2\nu+1)}) \\ &- 2\,(\mathsf{V}_{(\mathfrak{a})a}^{(2\nu+q)} - \mathsf{V}_{(\mathfrak{a})a}^{(2\nu+q+1)}) + (\mathsf{V}_{(\mathfrak{a})\mathfrak{a}}^{(2\nu+2q)} - \mathsf{V}_{(\mathfrak{a})\mathfrak{a}}^{(2\nu+2q+1)})\big], \end{aligned}$$

und daraus folgt nach (23), wenn wir bekannte Regeln über absolute Beträge anwenden:

$$(24)\qquad \left.\begin{matrix} L^{(2\nu)}(\gamma_{(i)} - \gamma^{(q)}_{(\mathfrak{i})}) \\ L^{(2\nu)}(\gamma_{(a)} - \gamma^{(q)}_{(\mathfrak{a})}) \end{matrix}\right\} \leqq 4\pi \cdot 8 \cdot 4\pi \mathsf{B}_g \eta_{2\nu} < (4\pi)^2 \cdot 9 \mathsf{B}_g \eta_{2\nu}.$$

Die Größen links konvergieren also mit wachsendem ν *gleichmäßig für alle Lagen von i und* $\mathfrak{i}$ *in* $\mathfrak{J}+\sigma$ *bzw. von a und $\mathfrak{a}$ in* $\mathfrak{A}+\sigma$ (unabhängig von q) gegen 0.

Nach diesen Feststellungen wenden wir uns jetzt den *allgemeinen Robinschen Potentialen* V, V', V'' usw. zu, die von *beliebigen* (stetigen) Randwerten g_s ausgehend gebildet zu denken sind [vgl. das Schema $(\mathfrak{R}_0)$]. Wir bilden mit ihnen das Integral

$$M^{(n)}(g, \gamma_{(o)} - \gamma^{(q)}_{(\mathfrak{o})}) = \mathsf{M}^{(n)}_{(o)}(g) - \mathsf{M}^{(n+q)}_{(\mathfrak{o})}(g) \qquad \text{[vgl. (4) und (6)]}.$$

Dieses hat nach (7i) bzw. (7a) den Wert

$$\pm 4\pi\Big((V_o^{(n)} \pm V_o^{(n+1)}) - (V_{\mathfrak{o}}^{(q+1)} \pm V_{\mathfrak{o}}^{(n+q+1)})\Big),$$

wo das obere oder untere Vorzeichen gilt, je nachdem die Punkte o und $\mathfrak{o}$ innerhalb oder außerhalb σ liegen. Die Schwarzsche Ungleichung $\operatorname{abs} M^{(n)}(g, \mathfrak{g}) \leqq \sqrt{L''(g) \cdot L^{(2n-2)}(\mathfrak{g})}$ [11]) liefert uns also unter Rücksicht auf (24) die Abschätzungen:

$$\left.\begin{matrix} \operatorname{abs}\Big((V_i^{(n)} + V_i^{(n+1)}) - (V_{\mathfrak{i}}^{(n+q)} + V_{\mathfrak{i}}^{(n+q+1)})\Big) \\ \operatorname{abs}\Big((V_a^{(n)} - V_a^{(n+1)}) - (V_{\mathfrak{a}}^{(n+q)} - V_{\mathfrak{a}}^{(n+q+1)})\Big) \end{matrix}\right\} \leqq \sqrt{L''(g)} \cdot \sqrt{9 \mathsf{B}_g \eta_{2n-2}}.$$

Für *Randpunkte* s und $\mathfrak{s}$ auf σ, für die beide Abschätzungen gleichzeitig zutreffen, gilt daher weiter

$$(25)\qquad \operatorname{abs}(V_s^{(n)} - V_{\mathfrak{s}}^{(n+q)}) < 3\sqrt{\mathsf{B}_g}\sqrt{L''(g)} \cdot \sqrt{\eta_{2n-2}},$$

und wegen des letzten Faktors [vgl. (22′)] folgt daher zunächst

$$(25')\qquad \lim_{n\to\infty} (V_s^{(n)} - V_{\mathfrak{s}}^{(n+q)}) = 0 \quad \text{gleichmäßig für alle } s \text{ und } \mathfrak{s} \text{ auf } \sigma \text{ (bei beliebigem } q).$$

Nun liegen aber bei Potentialen von Belegungen der Fläche σ sowohl der größte wie der kleinste aller *innerhalb* oder auf σ vorkommenden Werte am Rande; daher überträgt sich dieses letzte Resultat (25′) sofort von Randpunkten auch auf beliebige *innere* Punkte i und $\mathfrak{i}$:

$$\lim_{n\to\infty} (V_i^{(n)} - V_{\mathfrak{i}}^{(n+q)}) = 0 \quad \text{gleichmäßig für alle } i \text{ und } \mathfrak{i} \text{ (bei beliebigem } q).$$

---

[11]) An sich könnten wir unter gleichzeitiger Erhöhung des zweiten Index den ersten auf 0 herabsetzen, also L statt L'' schreiben, da auch bereits das Integral $L(g)$ stets einen bestimmten Sinn hat [Stud. S. 69 u. 73]. Wegen der Anwendungen in § 4 dürfte aber unser obiges Vorgehen den Vorzug verdienen.

Diese Formel ist aber der analytische Ausdruck für die Tatsache, daß *sich die Robinschen Potentiale* $V^{(n)}$ *im ganzen Gebiete* $\mathfrak{J}+\sigma$ *gleichmäßig einem konstanten Werte nähern*:

$$\lim_{n\to\infty} V_i^{(n)} = C \quad \text{gleichmäßig in } \mathfrak{J}+\sigma. \tag{26i}$$

Auf das *Außengebiet* $\mathfrak{A}$ kann man aus (25′) nicht ohne weiteres entsprechende Rückschlüsse ziehen. Da müssen wir uns vielmehr auf den Fall $\mathfrak{s} = s$ (und später $\mathfrak{a} = a$) beschränken. Dann allerdings können wir die Differenz $V_s^{(n)} - V_s^{(n+q)}$ ansehen als die Randwerte des Potentials einer Flächenbelegung *von der Gesamtmasse* 0 [vgl. $(\mathfrak{R}_0')$ S. 452], und erst bei diesen sind wir dann wieder sicher, daß *beide* Extremwerte (in $\mathfrak{A}+\sigma$) auf dem Rande σ liegen, und daß daher aus (25′) auch

$$\lim_{n\to\infty} (V_a^{(n)} - V_a^{(n+q)}) = 0 \quad \text{gleichmäßig für alle } a \text{ (bei beliebigem } q)$$

folgt, oder, daß *sich die Potentiale* $V^{(n)}$ *gleichmäßig im ganzen Gebiete* $\mathfrak{A}+\sigma$ *einer bestimmten Konvergenzfunktion nähern*:

$$\lim_{n\to\infty} V_a^{(n)} = F_a \quad \text{gleichmäßig in } \mathfrak{A}+\sigma. \tag{26a}$$

Von dieser in (26i) und (26a) auftretenden Konvergenzkonstanten C bzw. Konvergenzfunktion F_a (die wegen der Gleichmäßigkeit der Konvergenz bei σ stetig ineinander übergehen müssen) läßt sich dann leicht beweisen [vgl. Stud. S. 105—112 u. Beitr. S. 172], daß sie von der Natur der ursprünglich vorgeschriebenen Randwerte g_s nur insofern abhängen, als sie proportional mit der Masse $m = \frac{1}{2\pi}\int g_s\,d\sigma$ der Robinschen Flächenbelegungen sind:

$$C = m\cdot \Gamma_g, \qquad F_a = m\cdot \Pi_a, \tag{26′}$$

während der andere Faktor die schon oben eingeführte geometrische Konstante Γ_g bzw. eine rein geometrische, nur von der Natur der Fläche σ abhängige Funktion Π_a ist.

Durch Zusammenfassung der Formeln (25) und (26) ergibt sich dann noch bezüglich des *Grades der Annäherung*:

$$\left.\begin{array}{l} \operatorname{abs}(V_i^{(n)} - m\cdot\Gamma_g) \\ \operatorname{abs}(V_a^{(n)} - m\cdot\Pi_a) \end{array}\right\} < 3\sqrt{\mathsf{B}_g}\sqrt{L''(g)}\sqrt{\eta_{2n-2}}, \tag{27}$$

und von der hier auftretenden Funktion Π_a, die bei σ stetig in Γ_g übergeht, läßt sich leicht noch zeigen [vgl. Stud. S. 111], daß sie alle Eigenschaften eines Potentials aufweist. — *Es nähern sich also die Potentiale* $V^{(n)}$ *im ganzen Raume* $\mathfrak{J}+\mathfrak{A}$ *einer Funktion, die alle Eigenschaften des elektrostatischen Potentials besitzt*, d. h. des Potentials einer auf σ im

Gleichgewicht befindlichen elektrischen Ladung von der Masse m — nur eben das eine, daß sich diese Funktion wirklich als Potential einer auf σ ausgebreiteten Belegung darstellen läßt — das ergibt sich hier vorläufig nicht.

Die Formeln (27) lassen wieder den geometrischen Charakter der Konvergenz deutlich hervortreten [vgl. (22)].

§ 3.

Über die natürliche Belegung einer geschlossenen Fläche.

Daß jene Konvergenzfunktion $m \cdot \Pi_s$ nun tatsächlich das Potential einer einfachen Belegung der Fläche σ ist, oder, was dasselbe ist, daß sie bestimmte normale Ableitungen besitzt, oder auch: daß nicht nur die Potentiale V, V', V'' usw. einer bestimmten Funktion zustreben, sondern auch die Dichtigkeiten der aufeinanderfolgenden Robinschen Belegungen, oder die Flächenfunktionen g, g', g'' usw. sich einer Grenzfunktion nähern — diesen Nachweis habe ich im Anhang zu meinen „Beiträgen" geführt (S. 172—188) —, dort zwar nur im Gebiete des logarithmischen Potentials in der Ebene, doch gelten alle Schlüsse fast unverändert auch bei räumlichen Betrachtungen. — Diesen Beweis will ich nun hier nochmals in seinen wesentlichen Zügen auseinandersetzen, ihn dabei aber sogleich wieder so modifizieren, daß der geometrische Charakter der Konvergenz deutlich erkennbar wird.

Es handelt sich also darum, aus dem Verhalten der Werte der *Potentiale* V, V', V'' usw. selber Rückschlüsse zu ziehen auf ihre *normalen Ableitungen*. Zu diesem Zwecke bediene ich mich einer Methode, die man zutreffend etwa als die Methode der berührenden Kreise oder, hier im Raume, der berührenden Kugeln bezeichnen könnte. Um nämlich an einer Stelle s der im wesentlichen willkürlichen Fläche σ die normalen Ableitungen zu untersuchen, denke ich mir zwei die Fläche σ daselbst außen und innen berührende Kugeln konstruiert und untersuche, wie sich die radialen Ableitungen der *durch ihre Werte auf diesen Kugeln* (mittelst des Poissonschen Integrales) *ausgedrückten Potentiale* V, V', V'' usw. bei Annäherung an s verhalten.

Nun aber genügt bekanntlich *allein die Stetigkeit* der Potentialwerte auf σ *noch nicht* bei jenem Prozesse, die *Existenz* von Grenzwerten, eben der zu untersuchenden normalen Ableitungen, zu verbürgen; es müssen die Werte auf σ vielmehr noch *verschärften Stetigkeitsbedingungen* (ähnlich den bekannten Lipschitzschen Bedingungen) genügen. — Eine solche hinreichende Bedingung ist die folgende: Wenn es sich um die normalen Ableitungen des Potentials U einer einfachen Flächenbelegung von σ

$$U_p = \int \varkappa_\sigma \frac{d\sigma}{E_p} \qquad (\varkappa \text{ stetig auf } \sigma)$$

in einem Punkte s der Fläche σ handelt, konstruiere man um die in s errichtete Flächennormale einen kleinen Kreiszylinder vom Radius r und betrachte die Randwerte von U, die auf der Durchdringungskurve zwischen σ und diesem Zylinder anzutreffen sind, als Funktion des Azimutes ϑ der Zylindermeridianebenen und bilde unter Zugrundelegung dieser Auffassung

das „peripherische Mittel $\frac{1}{2\pi}\int\limits_{(r)} U\,d\vartheta$."

Dieses wird mit abnehmendem r, da U stetig ist, gegen den Wert U_s konvergieren. Setzen wir aber voraus, daß diese Konvergenz sogar rascher als die von $r^{\frac{3}{2}}$ gegen 0 vor sich geht, daß also eine Beziehung von der Form

$$\operatorname{abs}\left\{\frac{1}{2\pi}\int\limits_{(r)} U\,d\vartheta - U_s\right\} \leqq A\cdot r^{\frac{3}{2}}$$

besteht[12]), so sagen wir, die Randwerte von U erfüllten die *„peripherische Bedingung"*, und diese *ist nach Liapounoff hinreichend für die Existenz der normalen Ableitungen* — ja aber weiter habe ich (eben mittelst der Methode der berührenden Kugeln) bewiesen, daß sich, wenn diese Bedingung erfüllt ist, diese normalen Ableitungen *abschätzen* lassen durch die Schwankung $\Delta(U)$ der Funktion U[13]), allerdings nicht durchweg durch ihre erste, aber mindestens durch ihre $\frac{1}{4}$-te Potenz. Es gilt dann nämlich, gleichgültig ob wir die normale Ableitung außen oder innen bilden, die Abschätzung

$$(1)\qquad \operatorname{abs}\left(\frac{\partial U}{\partial \nu_s}\right) \leqq \mathfrak{a}_g \Delta(U) + \mathfrak{b}_g\left(A + \mathfrak{k}_g \mathfrak{M}(\varkappa)\right)^{\frac{3}{4}}\cdot\left(\Delta(U)\right)^{\frac{1}{4}} \quad \text{[Beitr. S. 178 u. 181]},$$

wo das Symbol $\mathfrak{M}$, wie stets im folgenden den absolut größten Wert der betreffenden Funktion bedeutet. — Diese Formel bringt die an sich so einleuchtende Tatsache zum Ausdruck, daß mit zur 0 abnehmender Schwankung $\Delta(U)$ (wenn sich also U im ganzen Raume mehr und mehr der 0 nähert), auch die normalen Ableitungen von U sich der 0 nähern.

Von dieser Formel (1) wollen wir nun die Anwendung auf die Robinschen Potentiale V, V', V'' usw. machen. Dann müssen wir zunächst zusehen,

---

[12]) Es würde genügen, anstatt $r^{\frac{3}{2}}$ zu schreiben $r^{1+\delta}$, wenn δ eine *beliebig kleine*, aber positive Größe ist.

[13]) Zunächst bedeutet $\Delta(U)$ die Schwankung in einem Gebiete, wie es von der Gesamtheit aller benutzten Kugeln erfüllt wird, doch kann man hier bei Betrachtungen im Raume auch die Schwankung *im Gesamtraume* darunter verstehen — während in der Ebene wegen des Verhaltens logarithmischer Potentiale im Unendlichen tatsächlich eine Beschränkung auf ein kleineres Gebiet beibehalten werden muß [vgl. Beitr. S. 183].

ob deren Randwerte die peripherische Bedingung erfüllen: Nun stellen die *Randwerte* V_s, V'_s, V''_s usw. sukzessiver Robinscher Potentiale aufeinanderfolgende Neumannsche Flächenfunktionen dar [Stud. S. 66], sind also den Funktionen $f, f' f''$ usw. vergleichbar. Von diesen Funktionen ist aber bekannt, daß, wenn die erste nur stetig ist, die nächste bereits „regulär stetig" ist, d. h. die verschärfte (Lipschitzsche) Stetigkeitsbedingung

$$\text{abs}\,(f'_{s_1} - f'_{s_2}) = d \cdot \sqrt{\mathsf{E}_{12}}\ ^{14)} \quad \text{mit} \quad d = \mathfrak{d}_g \mathfrak{M}(f)$$

befriedigt [Stud. S. 34], und das genügt dann, für die wieder nächstfolgende Funktion f'' den Nachweis für die Erfüllung der peripherischen Bedingung zu erbringen:

$$\text{abs}\left\{\frac{1}{2\pi}\int f''\,d\vartheta - f''_s\right\} \leqq A \cdot \varrho^{\frac{1}{2}} \quad \text{mit} \quad A = \mathfrak{e}_g \cdot d = \mathfrak{c}_g \mathfrak{M}(f) \quad \text{[Beitr. S. 42].}$$

Wir vergleichen demnach die Randwerte von $V^{(n)}$ unseren Werten f_s, so daß also die Randwerte von $V^{(n+2)}$ den Werten f''_s oder den Werten U_s in unseren obigen Ausführungen entsprechen; dann folgt nach Formel (1):

$$\text{abs}\left(\frac{\partial V^{(n+2)}}{\partial \nu_s}\right) \leqq \left[\mathfrak{a}_g (\varDelta(V^{(n+2)}))^{\frac{1}{2}} + \mathfrak{b}_g \left(\mathfrak{c}_g \mathfrak{M}(V^{(n)}) + \frac{\mathfrak{f}_g}{2\pi}\mathfrak{M}(g^{(n+2)})\right)^{\frac{1}{2}}\right] (\varDelta(V^{(n+2)}))^{\frac{1}{2}}.$$

Da aber aus dem Bildungsgesetz der Potentiale V, V', V'' usw. [vgl. $(\mathfrak{R}_0)$ S. 452] leicht Abschätzungen von der Form folgen:

$$\mathfrak{M}(V^{(n+2)}) \leqq C_g \cdot \mathfrak{M}(g^{(n+2)}) \leqq C_g \mathfrak{q}_g^2 \mathfrak{M}(g^{(n)}),$$

und daher sicher (als Differenz zweier Werte)

$$\varDelta(V^{(n+2)}) \leqq 2\,C_g \mathfrak{q}_g^2 \mathfrak{M}(g^{(n)})$$

ist, so ergibt sich schließlich für die normalen Ableitungen die folgende Abschätzung:

$$\text{abs}\left(\frac{\partial V^{(n+2)}}{\partial \nu_s}\right) \leqq \left[\mathfrak{a}_g (2\,C_g \mathfrak{q}_g^2)^{\frac{1}{2}} + \mathfrak{b}_g \left(\mathfrak{c}_g C_g + \frac{\mathfrak{f}_g}{2\pi}\mathfrak{q}_g^2\right)^{\frac{1}{2}}\right] (\mathfrak{M}(g^{(n)}))^{\frac{1}{2}} (\varDelta(V^{(n+2)}))^{\frac{1}{2}},$$

oder bei Zusammenfassung aller geometrischen Konstanten in eine neue[15]):

$$\text{(2)} \qquad \text{abs}\left(\frac{\partial V^{(n+2)}}{\partial \nu_s}\right) \leqq \mathfrak{Q}_g (\mathfrak{M}(g^{(n)}))^{\frac{1}{2}} (\varDelta(V^{(n+2)}))^{\frac{1}{2}},$$

---

[14]) E_{12} ist dabei die Entfernung der Punkte s_1 und s_2. — An Stelle von $\sqrt{\mathsf{E}_{12}} = \mathsf{E}_{12}^{\frac{1}{2}}$ hätte man wieder eine beliebige noch niedrigere Potenz E_{12}^{δ} verwenden können [vgl. Stud. S. 32 und oben Anmerkung [12])].

[15]) Auf die Werte der durch die Marke gekennzeichneten geometrischen Konstanten kommt es im einzelnen überhaupt nicht an, sondern nur eben auf die Feststellung ihres rein geometrischen Charakters.

und zwar gilt diese Formel, wie schon bemerkt, in gleicher Weise für die innen wie die außen gebildete normale Ableitung, d. h. nach dem Schema $(\mathfrak{N}_0)$ in gleicher Weise für $g_s^{(n+3)} - g_s^{(n+2)}$ wie für $g_s^{(n+3)} + g^{(n+2)}$ und daher auch für die halbe Differenz, d. h. für $g_s^{(n+2)}$, und da überdies die Abschätzung noch von der Wahl des Punktes s auf σ unabhängig ist, bleibt sie auch noch für den absolut größten Wert von $g_s^{(n+2)}$ gültig, so daß wir erhalten:

$$(3) \qquad \mathfrak{M}^4(g^{(n+2)}) \leq \mathfrak{Q}_g^4\, \Delta(V^{(n+2)})\, \mathfrak{M}^3(g^{(n)}).$$

Diese Beziehung zwischen den absolut größten Werten der verschiedenen Robinschen Flächenfunktionen $g^{(n)}$ bildet nun die *Grundlage der weiteren Untersuchung.*

Zunächst folgt bezüglich des Faktors $\Delta(V^{(n+2)})$ nach dem Schlußergebnis von § 2, daß sich diese Schwankung mehr und mehr der von $m \cdot \Pi$, d. i. dem Werte $|m| \cdot \Gamma_g$ nähern muß, und zwar ergibt sich nach (27) S. 464 genauer

$$(4) \qquad \begin{aligned} \Delta(V^{(n+2)}) &< |m| \cdot \Gamma_g + 6\sqrt{\mathrm{B}_g}\sqrt{L''(g)}\sqrt{\eta_{2n+2}} \\ &< |m| \cdot \Gamma_g + 6\sqrt{\mathrm{B}_g}\sqrt{L''(g)}\sqrt{\eta_2} \qquad (\text{weil } \eta_{2n+2} \leq \eta_2 \text{ ist}). \end{aligned}$$

Nunmehr führen wir noch den absolut größten Wert der vorgegebenen Randwertfunktion g,

$$\mathfrak{M}(g) = \mu$$

ein. Dann ist augenscheinlich

$$(5\text{a}) \qquad |m| \leq \int \mu\, d\sigma = \frac{\sigma_g}{2\pi} \cdot \mu,$$

wenn σ_g die Größe der Fläche σ, ihr Areal, bedeutet. Ferner ist

$$L''(g) = [V, V'']_{\mathfrak{T}} = 2\int V_\sigma g_\sigma''\, d\sigma \qquad \text{[Stud. S. 73]}$$

und daraus folgt unter Benutzung schon oben eingeführter geometrischer Konstanten:

$$(5\text{b}) \qquad L''(g) \leq 2\int (C_g \cdot \mu)(\mathrm{q}_g^2 \mu)\, d\sigma = 2 C_g \sigma_g \mathrm{q}_g^2 \mu^2.$$

Diese beiden Abschätzungen (5a) und (5b) berücksichtigen wir in (4), um zu erhalten

$$\Delta(V^{(n+2)}) < \mathfrak{C}_g \cdot \mu \qquad \left(\mathfrak{C}_g = \frac{\sigma_g}{2\pi}\Gamma_g + \mathrm{q}_g\sqrt{2 C_g \sigma_g}\right),$$

und Formel (3) nimmt daher die Gestalt an:

$$(6) \qquad \mathfrak{M}^4(g^{(n+2)}) < \mathfrak{F}_g \cdot \mu\, \mathfrak{M}^3(g^{(n)}) \qquad (\mathfrak{F}_g = \mathfrak{Q}_g^4 \cdot \mathfrak{C}_g).$$

Machen wir nun, um uns zunächst auf gerade Indizes zu beschränken, die Annahme, es sei einmal

$$\mathfrak{M}(g^{(2\nu)}) \geq 2\mathfrak{F}_g \cdot \mu, \quad \text{so ist} \quad \left(\frac{\mathfrak{M}(g^{(2\nu+2)})}{\mathfrak{M}(g^{(2\nu)})}\right)^4 < \frac{\mathfrak{F}_g \cdot \mu}{2\mathfrak{F}_g \cdot \mu} = \frac{1}{2},$$

d. h. es nehmen die Größen $\mathfrak{M}(g^{(2\nu)})$ rascher ab als die Glieder einer geometrischen Reihe mit dem Quotienten $\sqrt[4]{\frac{1}{2}}$; nach einer gewissen Anzahl von Schritten muß also ein $\mathfrak{M}(g^{(2\nu)}) < 2\mathfrak{J}_g \cdot \mu$ werden, und dann lehrt (6) sofort, daß auch alle folgenden Größen $< 2\mathfrak{J}_g \cdot \mu$ bleiben. Wenn also überhaupt einige der Größen $\mathfrak{M}(g^{(2\nu)}) > 2\mathfrak{J}_g \cdot \mu$ werden, so können es nur die ersten sein, und diese nehmen dauernd ab, sind also sämtlich kleiner als die allererste, d. h. als $\mathfrak{M}(g) = \mu$ — oder mit anderen Worten: Die Maxima der Funktionen $g^{(n)}$ mit geradem Index liegen sämtlich unterhalb der größeren der beiden Zahlen $2\mathfrak{J}_g \cdot \mu$ und μ, und ebenso die Größen $\mathfrak{M}(g^{(2\nu+1)})$ unterhalb der größeren der Zahlen $2\mathfrak{J}_g \cdot \mu$ und $\mathfrak{M}(g') \leqq \mathfrak{q}_g \cdot \mu$ — für *sämtliche* $\mathfrak{M}(g^{(n)})$ gilt daher sicher

$$(7) \qquad \mathfrak{M}(g^{(n)}) < \mathfrak{K}_g \mu, \quad \text{wenn} \quad \mathfrak{K}_g = 1 + \mathfrak{q}_g + 2\mathfrak{J}_g.$$

Damit ist also zunächst bewiesen, daß *sämtliche Robinschen Flächenfunktionen $g^{(n)}$ unterhalb einer endlichen Grenze bleiben.*

Nunmehr wenden wir uns zu dem

Spezialfall: $\qquad m = \frac{1}{2\pi}\int g_\sigma \, d\sigma = 0$

und greifen in ihm zur Abschätzung von $\Delta(V^{(n+2)})$ in (3) auf die erste Ungleichung (4) zurück. So erhalten wir

$$(3_0) \qquad \mathfrak{M}^4(g^{(n+2)}) < \mathfrak{Q}_g^4 \, 6 \sqrt{\mathsf{B}_g} \sqrt{L''(g)} \sqrt{\eta_{2n+2}} \, \mathfrak{M}^3(g^{(n)})$$

und weiter nach (5b) und (7):

$$\mathfrak{M}^4(g^{(2n+2)}) < 6\mathfrak{Q}_g^4 \sqrt{2\mathsf{B}_g C_g \sigma_g \mathfrak{q}_g \mu \cdot \mathfrak{K}_g^3 \mu^3} \sqrt{\eta_{2n+2}},$$

oder schließlich das Resultat:

$$(8_0) \qquad \mathfrak{M}(g^{(n)}) < \mathfrak{L}_g \mu \sqrt[8]{\eta_{2n-2}} \qquad (n \geqq 2),$$

das bei Berücksichtigung der Eigenschaft (22′) der geometrischen Größen η_n (S. 462) besagt, daß *in dem Falle* $m = 0$ *die sämtlichen Funktionen* $g^{(n)}$ *gleichmäßig gegen* 0 *konvergieren*:

$$(9_0) \qquad \lim_{n \to \infty} g_s^{(n)} = 0 \quad \textit{falls} \quad m = 0 \qquad \text{gleichmäßig für alle } s \text{ auf } \sigma.$$

Nach dieser Feststellung nehmen wir wieder den

allgemeinen Fall: $\qquad$ *m beliebig*,

auf, betrachten als Ausgangsfunktion jetzt aber nicht g selber, sondern $g - g^{(q)}$ (q beliebig $= 1, 2, 3, \ldots$); dann ist

$$\frac{1}{2\pi}\int (g_\sigma - g_\sigma^{(q)}) \, d\sigma = 0 \qquad [\text{vgl. } (\mathfrak{R}_0') \text{ S. } 452],$$

und es gelten daher für die von diesen Werten aus gebildeten Robinschen Flächenfunktionen $g^{(n)} - g^{(n+q)}$ die oben im Spezialfalle abgeleiteten Resultate. Nun ist aber nach (7)

$$\mathfrak{M}(g^{(n)} - g^{(n+q)}) \leqq \mathfrak{M}(g^{(n)}) + \mathfrak{M}(g^{(n+q)}) < 2\,\mathfrak{K}_g\,\mu$$

und ferner nach (2r) und (3) S. 453:

$$L''(g - g^{(q)}) \leqq L''(g - g^{(q)}) + L''(g + g^{(q)}) = 2\,L''(g) + 2\,L^{(2q+2)}(g) \leqq 4\,L''(g).$$

Demgemäß können wir jetzt (bei $g - g^{(q)}$ als Ausgangsfunktion) in (3_0) an Stelle der früher beim Übergang zu (8_0) als obere Grenzen von $\mathfrak{M}(g^{(n)})$ und $\sqrt{L''(g)}$ benutzten Werte (7) und (5b) immer genau das Doppelte einrücken, und das liefert dann, wie man leicht übersieht, anstatt (8_0) das Resultat

$$(8) \qquad \mathfrak{M}(g^{(n)} - g^{(n+q)}) < 2\,\mathfrak{L}_g\,\mu \sqrt[8]{\eta_{2n-2}}$$

und hieraus folgt dann wieder, daß man durch Vergrößerung allein von n (unabhängig von q) die Differenzen $g_s^{(n)} - g_s^{(n+q)}$ gleichmäßig für alle Punkte s von σ unter jeden Kleinheitsgrad herabdrücken kann, d. h. aber: *im Falle $m \neq 0$ nähern sich die Funktionen $g^{(n)}$ gleichmäßig einer bestimmten Konvergenzfunktion*, die wir gleich $2\pi m \cdot \varrho$ setzen:

$$(9) \qquad \lim_{n \to \infty} g_s^{(n)} = 2\pi m \cdot \varrho_s \qquad \text{gleichmäßig für alle } s \text{ auf } \sigma.$$

Dann folgt leicht, eben infolge der Gleichmäßigkeit der Konvergenz (durch Grenzübergang aus den Formeln $(\mathfrak{R}_0')$ und (26i)):

$$\int \varrho_s\, d\sigma = 1 \quad \text{und} \quad \int \frac{\varrho_s\, d\sigma}{E_t} = \Gamma_g \qquad \text{[vgl. Beitr. S. 187]}$$

und damit das Resultat, daß ϱ die Dichtigkeit der sogenannten *natürlichen Belegung* ist, d. h. einer Belegung von der Gesamtmasse 1, deren Potential in $\mathfrak{J} + \sigma$ konstant ist. — Zugleich ist damit aber *der Beweis für die Existenz der natürlichen Belegung erbracht.*

Im Endresultat lassen sich die beiden bisher gesondert behandelten Fälle $m = 0$ und $m \neq 0$ leicht folgendermaßen zusammenfassen:

Die nach dem Schema $(\mathfrak{R}_0)$ S. 452, ausgehend von beliebigen stetigen Randwerten g_s sukzessive gebildeten Robinschen Flächenfunktionen g, g', g'' usw. nähern sich gleichmäßig in allen Punkten s von σ den Werten $2\pi m \cdot \varrho_s$,

$$\lim_{n \to \infty} g_s^{(n)} = 2\pi m \cdot \varrho_s \quad \textit{mit} \quad m = \frac{1}{2\pi} \int g_s\, d\sigma,$$

wenn ϱ die Dichtigkeit der natürlichen Belegung der Fläche σ bedeutet.

Besonders hervorgehoben sei wieder der geometrische Charakter der Konvergenz. Sie wird wieder erzwungen durch das Verhalten rein

geometrischer Größen, nämlich der Zahlen η_2, η_3, η_4 usw., während die spezielle Natur der Randwerte g_s bei den Abschätzungen (8_0) und (8) nur zum Ausdruck kommt in einem sich stets gleichbleibendem Faktor, hier in μ, dem absolut größten der Werte g_s.

§ 4.

Konfigurationskonstanten höherer Ordnung und der schließliche Konvergenzbeweis für die Neumannschen und Robinschen Reihen.

Beschäftigten wir uns bisher mit den „Konvergenzeigenschaften" der Neumannschen und Robinschen Flächenfunktionen und Potentiale [vgl. die Schemata $(\mathfrak{N}_0)$ und $(\mathfrak{R}_0)$], d. h. stellten wir fest, was bei wachsendem Index aus den einzelnen Funktionen (oder wenigstens einfachen Verbindungen von ihnen wie $f_s^{(n)} - f_s^{(n+2)}$, $W_i^{(n)} - W_i^{(n+1)}$, $W_a^{(n)} + W_a^{(n+1)}$) wird, also aus den einzelnen Elementen der Neumannschen und Robinschen Reihen $(\mathfrak{N}_1)$ und $(\mathfrak{R}_1)$, so soll jetzt die Konvergenz dieser Reihen selbst untersucht werden — zunächst der Neumannschen Reihen.

Unter Benutzung der polaren Flächenfunktionen $\gamma_{(o)}^{(n)}$ [vgl. (8) S. 454] habe ich die Neumannschen Potentiale $W^{(n)}$ folgendermaßen darzustellen gelehrt:

$$(10) \qquad W_o^{(n)} = \frac{1}{2\pi}\int f_\sigma \gamma_{(o)\sigma}^{(n)}\, d\sigma \qquad \text{[Stud. S. 94].}$$

Es ist dies eine Formel, die man als Ausdruck einer Umkehr der Integrationsfolge ansehen kann: Während nach der ursprünglichen Definition durch das Schema $(\mathfrak{N}_0)$ die Potentiale $W^{(n)}$ $(n+1)$-fache Integrale sind, bei denen die Ausgangsrandwerte f_s unter dem innersten Integrale stehen, treten diese Werte hier unter einem einfachen, letzten Integrale auf, während die übrigen n Integrationen in dem anderen rein geometrischen Faktor unter diesem Integral enthalten sind.

Diese letzte Formel wenden wir nun zunächst auf den Fall an, daß der Punkt o nacheinander zwei beliebige innere Lagen (i bzw. $\mathfrak{i}$) annimmt. Dann ist also (wenn wir noch $q-1$ statt n schreiben):

$$(11\mathrm{i}) \qquad W_i^{(q-1)} - W_{\mathfrak{i}}^{(q-1)} = \frac{1}{2\pi}\int f_\sigma \left[\gamma_{(i)\sigma}^{(q-1)} - \gamma_{(\mathfrak{i})\sigma}^{(q-1)}\right] d\sigma,$$

und entsprechend führt die Anwendung auf zwei äußere Punkte a und $\mathfrak{a}$ zu der Formel

$$(11\mathrm{a}) \qquad W_a^{(q-1)} - W_{\mathfrak{a}}^{(q-1)} = \frac{1}{2\pi}\int f_\sigma \left[\gamma_{(a)\sigma}^{(q-1)} - \gamma_{(\mathfrak{a})\sigma}^{(q-1)}\right] d\sigma.$$

Nun aber wollen wir von Betrachtungen im Innen- oder Außengebiet übergehen zu solchen auf der Fläche σ selbst. Wir lassen daher die beiden

Punkte i und a gegen einen Randpunkt s gehen und $\mathfrak{i}$ und $\mathfrak{a}$ gegen einen anderen $\mathfrak{s}$ und beachten dann einerseits, daß nach dem Schema $(\mathfrak{N}_0)$

$$\frac{1}{2}(W_{is}^{(q-1)} + W_{as}^{(q-1)}) = f_s^{(q)}$$

ist, und setzen andrerseits zur Abkürzung:

$$(12) \qquad \frac{1}{2}(\gamma_{(is)\sigma}^{(q-1)} + \gamma_{(as)\sigma}^{(q-1)}) = \gamma_{(s)\sigma}^{(q-1)}$$

— von $n = 2$ an bleiben die Polarfunktionen $\gamma_{(0)}^{(n)}$ auch bei Annäherung des Poles an σ stetig, und auch schon für $n = 0$ oder 1 werden sie höchstens integrabel unstetig —, dann ergibt sich durch Kombination von (11 i) und (11 a):

$$(11\,s) \qquad f_s^{(q)} - f_{\mathfrak{s}}^{(q)} = \frac{1}{2\pi}\int f_\sigma [\gamma_{(s)\sigma}^{(q-1)} - \gamma_{(\mathfrak{s})\sigma}^{(q-1)}]\, d\sigma$$

und, da sich nach dem Schlußresultat von § 3 beide Funktionen $\gamma_{(s)}^{(q-1)}$ und $\gamma_{(\mathfrak{s})}^{(q-1)}$ derselben Konvergenzfunktion $2\pi \cdot 1 \cdot \varrho$ nähern [vgl. unten (12′)], so ersieht man aus dieser Formel bereits, daß mit wachsendem q die Differenzen $f_s^{(q)} - f_{\mathfrak{s}}^{(q)}$ gegen 0 gehen werden. Es ließe sich auch leicht zeigen, daß diese Konvergenz sogar gleichmäßig für alle Lagen von s und $\mathfrak{s}$ auf σ erfolgt.

Um aber weiterzugehen, zerlegen wir jetzt die Fläche σ in zwei Teile α und β, indem wir jedes Element $d\sigma$ zum Teile α oder aber β rechnen, je nachdem in ihm $\gamma_{(s)}^{(q-1)}$ größer oder kleiner als $\gamma_{(\mathfrak{s})}^{(q-1)}$ ist. Dann folgt

$$(13) \quad f_s^{(q)} - f_{\mathfrak{s}}^{(q)} \leqq G \cdot \frac{1}{2\pi}\int_\alpha [\gamma_{(s)\alpha}^{(q-1)} - \gamma_{(\mathfrak{s})\alpha}^{(q-1)}]\, d\alpha + K \cdot \frac{1}{2\pi}\int_\beta [\gamma_{(s)\beta}^{(q-1)} - \gamma_{(\mathfrak{s})\beta}^{(q-1)}]\, d\beta,$$

wenn G und K den größten und kleinsten Wert von f bedeuten.

Nun ist nach bekanntem Gaußschen Satz

$$(14) \qquad \int \gamma_{(i)\sigma}\, d\sigma = 4\pi \quad \text{und} \quad \int \gamma_{(a)\sigma}\, d\sigma = 0 \qquad \text{[vgl. (8) S. 454]}$$

für beliebige Lagen von i in $\mathfrak{J} + \sigma$ und von a in $\mathfrak{A} + \sigma$ und also wegen der Massengleichheit aufeinanderfolgender Robinscher Flächenbelegungen [vgl. $(\mathfrak{N}_0')$] auch

$$\int \gamma_{(i)\sigma}^{(q-1)}\, d\sigma = 4\pi, \quad \int \gamma_{(a)\sigma}^{(q-1)}\, d\sigma = 0,$$

und daraus folgt nach (12):

$$(12') \qquad \int \gamma_{(s)\sigma}^{(q-1)}\, d\sigma = 2\pi \qquad \text{(für alle } s \text{ auf } \sigma)$$

und demgemäß ist die Summe der in (13) auftretenden Integrale

$$\int_\alpha + \int_\beta = \int_\sigma [\gamma_{(s)\sigma}^{(q-1)} - \gamma_{(\mathfrak{s})\sigma}^{(q-1)}]\, d\sigma = 0,$$

oder aber es folgt:

$$\int_\alpha = -\int_\beta = \frac{1}{2}\int_\sigma \operatorname{abs}\left[\gamma^{(q-1)}_{(s)\sigma} - \gamma^{(q-1)}_{(\mathfrak{s})\sigma}\right] d\sigma$$

und demnach ergibt sich aus (13) weiter:

$$(15)\qquad f^{(q)}_s - f^{(q)}_{\mathfrak{s}} \leqq (G-K)\frac{1}{4\pi}\int \operatorname{abs}\left[\gamma^{(q-1)}_{(s)\sigma} - \gamma^{(q-1)}_{(\mathfrak{s})\sigma}\right] d\sigma.$$

Der hier neben $G-K$ stehende Faktor ist eine außer von den geometrischen Verhältnissen lediglich von der Lage der Punkte s und $\mathfrak{s}$ auf σ abhängige Größe (die man in Analogie mit einer früher von mir angewandten Bezeichnungsweise etwa die Öffnungsfunktion q-ter Ordnung $D^{(q)}(s, \mathfrak{s})$ der Fläche σ in bezug auf die Pole s und $\mathfrak{s}$ nennen könnte). — Das *Maximum* dieser Größe für alle möglichen Lagen von s und $\mathfrak{s}$ auf σ, das jetzt also eine *allein von den geometrischen Verhältnissen* abhängige Zahl ist,

$$(16)\qquad c^{(q)}_g = \operatorname{Max}\left(\frac{1}{4\pi}\int \operatorname{abs}\left[\gamma^{(q-1)}_{(s)\sigma} - \gamma^{(q-1)}_{(\mathfrak{s})\sigma}\right] d\sigma\right) \quad (= \operatorname{Max} D^{(q)}(s, \mathfrak{s}))$$

heiße die „*Konfigurationskonstante q-ter Ordnung der Fläche σ*". Mit ihrer Hilfe kann man dann sogar die *größte* Wertdifferenz von $f^{(q)}$, also die Schwankung $\Delta(f^{(q)})$ abschätzen:

$$\Delta(f^{(q)}) \leqq (G-K)\cdot c^{(q)}_g.$$

Da nun aber hierbei die Ausgangsfunktion f mit der Schwankung $\Delta(f) = G-K$ ganz beliebig (nur stetig) vorausgesetzt war, können wir sie auch ersetzen durch eine beliebige spätere Funktion $f^{(n)}$ und erhalten so:

$$(17)\qquad \Delta(f^{(n+q)}) \leqq \Delta(f^{(n)})\cdot c^{(q)}_g.$$

Die Konfigurationskonstante q-ter Ordnung stellt also die (kleinste) obere Grenze für den Quotienten zwischen den Schwankungen einer beliebigen der Flächenfunktionen f, f', f'' usw. und der q-ten folgenden dieser Funktionen dar — gerade so wie die ursprüngliche C. Neumannsche Konfigurationskonstante c_g (erster Ordnung, also identisch mit unserem c'_g) die obere Grenze für den Quotienten der Schwankungen zweier *unmittelbar* aufeinanderfolgender solcher Funktionen darstellte [vgl. die Einleitung S. 447], so daß also diese Konfigurationskonstanten höherer Ordnung tatsächlich als ganz naturgemäße Verallgemeinerungen der ursprünglich von C. Neumann eingeführten Konstanten anzusehen sind.

Mit diesen höheren Konfigurationskonstanten wollen wir uns nun noch näher beschäftigen. Da $\int[\gamma_{(s)\sigma} - \gamma_{(\mathfrak{s})\sigma}]\, d\sigma = 0$ ist [vgl. (14)], so können wir die in § 3 für den „Spezialfall" ($m = 0$) angestellten Betrachtungen anwenden, wenn wir $g = \gamma_{(s)} - \gamma_{(\mathfrak{s})}$ setzen. Da dann allerdings g und g'

bei Annäherung von i oder $\mathfrak{i}$ an σ nicht mehr beschränkt bleiben, betrachten wir gleichsam als Ausgangswert erst die Werte $g_s'' = \gamma_{(i)s}'' - \gamma_{(\mathfrak{i})s}''$, so daß also, wenn wir $\mathfrak{M}(g^{(q-1)})$ abschätzen wollen, in Formel (8_0) $n = q - 3$ zu setzen ist, und sich daher als Abschätzung

$$\mathfrak{M}(\gamma_{(i)}^{(q-1)} - \gamma_{(\mathfrak{i})}^{(q-1)}) < \mathfrak{L}_g \bar{\mu} \sqrt[8]{\eta_{2q-8}} \qquad (q \geqq 5)$$

ergibt, wenn jetzt $\bar{\mu}$ den (sicher endlichen) absolut größten Wert bedeutet, den $\gamma_{(i)s}'' - \gamma_{(\mathfrak{i})s}''$ bei beliebigen Lagen von s auf σ und der Pole i und $\mathfrak{i}$ in $\mathfrak{J} + \sigma$ annimmt. — Eine ganz entsprechende Abschätzung erhält man für $\mathfrak{M}(\gamma_{(a)}^{(q-1)} - \gamma_{(\mathfrak{a})}^{(q-1)})$, nur daß an Stelle der allein noch von den geometrischen Verhältnissen der Fläche σ abhängigen Zahl $\bar{\mu}$ im allgemeinen eine andere, $\bar{\bar{\mu}}$, treten wird. Bezeichnen wir nun noch die größere dieser beiden Zahlen $\bar{\mu}$ und $\bar{\bar{\mu}}$ mit μ_g, so erhalten wir nach (12) für alle beliebigen Lagen von s und $\mathfrak{s}$ auf σ die Abschätzung

$$\mathfrak{M}(\gamma_{(s)}^{(q-1)} - \gamma_{(\mathfrak{s})}^{(q-1)}) < \mathfrak{L}_g \mu_g \sqrt[8]{\eta_{2q-8}},$$

und mit ihrer Hilfe weiter nach (16):

$$c_g^{(q)} < \frac{1}{4\pi} \mathfrak{L}_g \mu_g \sigma_g \sqrt[8]{\eta_{2q-8}} \qquad (18) \quad (q \geqq 5).$$

Mit dieser Abschätzung ist also nach der Grundeigenschaft (22′) der Größen η_n das Hauptresultat gewonnen: *die höheren Konfigurationskonstanten nehmen mit wachsender Ordnungszahl gegen* 0 *ab.*

Nach dieser wichtigen Feststellung wenden wir uns zu den eigentlichen Konvergenzbeweisen.

A. Die Neumannschen Reihen.

Aus unserem letzten Resultat folgt, daß man in der Reihe der Konfigurationskonstanten c_g', c_g'', c_g''' usw. stets nach einer gewissen Anzahl von Schritten zu einer solchen Konstanten kommt, die kleiner als 1 ist,

(19) q *existiert, so daß* $c_g^{(q)} < 1$.

Im folgenden wollen wir uns nun q stets in dieser Weise bestimmt und dann festgehalten denken (es sei etwa die kleinste Zahl mit dieser Eigenschaft). Dann kann man den Beweis für die Existenz der Konvergenzkonstanten C der Funktionen f, f', f'' usw. sowie für die Konvergenz der Neumannschen Reihen von Potentialen im wesentlichen nach C. Neumann selbst führen[16]),

[16]) Vgl. die Darstellung dieses Verfahrens z. B. in meiner Arbeit in Math. Annalen **55**.

der die letztere Frage zurückführt auf die nach der Konvergenz der folgenden *Reihen von Oberflächenfunktionen*:

$$(20) \qquad \begin{aligned} &(f-f')+(f''-f''')+(f^{\mathrm{IV}}-f^{\mathrm{V}})+\cdots, \\ &(C-f)+(C-f')\;+(C-f'')\;+\cdots. \end{aligned}$$

Zum Beweise der (gleichmäßigen) Konvergenz dieser Reihen braucht man nämlich jetzt immer nur q Glieder zusammenzufassen, um wieder auf eine Abschätzung durch eine geometrische Reihe mit echtgebrochenen Quotienten $c_q^{(q)}$ zu kommen, oder man kann auch sagen: Die Reihen konvergieren mindestens ebenso rasch wie eine geometrische Reihe mit dem Quotienten $\sqrt[q]{c_q^{(q)}}$. — Der Beweis ist so leicht zu übersehen, daß sich ein näheres Eingehen auf ihn erübrigt.

Somit ist also der Beweis für die Anwendbarkeit der Methode des arithmetischen Mittels auf Gebiete, deren Grenzflächen nur sehr allgemeinen Bedingungen entsprechen, geführt — ein Beweis, der schließlich in engster Anlehnung an den ursprünglichen C. Neumannschen Gedankengang mit der Konfigurationskonstanten verläuft: C. Neumann führt den Beweis nur für den Fall, daß seine Konfigurationskonstante (erster Ordnung) kleiner als 1 ist — wir haben hier bewiesen, daß, auch wenn das nicht der Fall ist, doch *sicher eine Konfigurationskonstante höherer Ordnung kleiner als* 1 wird, und daß auch das genügt, die Konvergenz zu verbürgen.

Es führen diese Überlegungen dazu, die geschlossenen Flächen σ in verschiedene Klassen zu teilen nach der Ordnung der niedrigsten echtgebrochenen Konfigurationskonstanten, und man könnte dann etwa sagen, die Konvergenz der Methode des arithmetischen Mittels für diese verschiedenen Klassen sei von der ersten, zweiten usw. Ordnung. — Unter Benutzung dieser Ausdrucksweise ließe sich dann unser Resultat mit dem in der Einleitung erwähnten C. Neumannschen so zusammenfassen: *Für alle geschlossenen konvexen Flächen konvergiert die Methode des arithmetischen Mittels von der ersten Ordnung, aber auch für jede andere geschlossene Fläche (mit den von uns immer gemachten Einschränkungen) findet Konvergenz von einer endlichen Ordnung statt.*

B. Die Robinschen Reihen.

Nach einem früher von mir bewiesenen und schon oben einmal erwähnten Resultate [Stud. S. 66] bilden die Oberflächenwerte V_s, V_s', V_s'' usw. der Robinschen Potentiale eine Folge aufeinanderfolgender Neumannscher Flächenfunktionen (wie f_s, f_s', f_s'' usw.), und aus der soeben allgemein bewiesenen gleichmäßigen Konvergenz der Reihen (20) folgt also auch die

der Reihen

$$(21) \quad \begin{array}{l} (V_s - V_s') + (V_s'' - V_s''') + (V_s^{\mathrm{IV}} - V_s^{\mathrm{V}}) + \cdots \\ V_s + V_s' + V_s'' + V_s''' + \cdots \end{array} \quad \begin{array}{r} (\text{falls } m = 0 \text{ und damit } C = m\Gamma_s = 0; \\ \text{vgl. } (26'), \text{ S. } 464). \end{array}$$

Dies sind aber im wesentlichen die Werte der Robinschen Reihen ($\mathfrak{R}_1$) S. 452, gebildet in Oberflächenpunkten. Da aber bei Potentialen von Flächenbelegungen unter den Oberflächenwerten die absolut größten *aller* Werte enthalten sind, so folgt aus der Konvergenz der Reihen (21) auch sofort, daß *die Robinschen Reihen gleichmäßig im ganzen Außen- und Innengebiete konvergieren* — denn die für die Konvergenz der zweiten Reihe (21) notwendige Voraussetzung, daß

$$m = \frac{1}{2\pi}\int g_s \, d\sigma = 0$$

ist, muß bei der zweiten Randwertaufgabe für ein Innengebiet bekanntlich stets erfüllt sein.

Marburg, den 11. Februar 1929.

(Eingegangen am 15. 2. 1929.)

La formule de H. Bruns et la théorie des figures planétaires.

Von

R. Wavre in Genf.

1. La formule de Bruns.

Considérons une masse fluide dont les différentes particules s'attirent suivant la loi de Newton, en équilibre relatif dans sa rotation autour d'un axe. Soient: ω la vitesse angulaire, ϱ la densité, p la pression, U le potentiel newtonien, oz l'axe de rotation et ox, oy, oz un trièdre trirectangle. L'hydrodynamique fournit les trois équations suivantes, comme on le sait,

$$\frac{1}{\varrho}\frac{\partial p}{\partial x}=\frac{\partial U}{\partial x}+\omega^2 x,\quad \frac{1}{\varrho}\frac{\partial p}{\partial y}=\frac{\partial U}{\partial y}+\omega^2 y,\quad \frac{1}{\varrho}\frac{\partial p}{\partial z}=\frac{\partial U}{\partial z}. \tag{1}$$

Jointes à la condition supplémentaire que la densité croisse de la surface vers le centre de l'astre, elles expriment la condition nécessaire et suffisante pour qu'il y ait équilibre relatif.

Les équations (1) impliquent une relation de la forme

$$\varrho = f(p). \tag{2}$$

Soient Φ le potentiel du champ de la pesanteur, Q celui de la force centrifuge:

$$\Phi=\int\frac{1}{\varrho}\frac{dp}{d\varrho}\,d\varrho, \tag{3}$$

$$Q=\frac{\omega^2}{2}(x^2+y^2). \tag{4}$$

Les trois équations (1) se résument par la relation entre les potentiels

$$\Phi = U + Q. \tag{5}$$

Si ∇ représente le symbole de Laplace, l'équation (5) implique la relation suivante

$$\nabla\Phi = \nabla U + \nabla Q. \tag{6}$$

La formule de Poisson et l'expression (4) de Q montrent que l'équation (6) peut s'écrire

$$\nabla \Phi = 2\omega^2 - 4\pi i \varrho \tag{6'}$$

où i représente la constante de l'attraction universelle. Soient g l'intensité de la pesanteur et α, β, γ les cosinus directeurs de la direction du fil à plomb par rapport aux axes x, y, z. On a, comme on sait,

$$\frac{\partial \Phi}{\partial x} = g\alpha, \quad \frac{\partial \Phi}{\partial y} = g\beta, \quad \frac{\partial \Phi}{\partial z} = g\gamma \tag{7}$$

et, on le vérifie facilement,

$$\nabla \Phi = \frac{dg}{dn}(\alpha^2 + \beta^2 + \gamma^2) + g\left(\frac{\partial \alpha}{\partial x} + \frac{\partial \beta}{\partial y} + \frac{\partial \gamma}{\partial z}\right)$$

dn désignant un élément de normale à la surface $\Phi =$ constante dirigé vers le bas.

Or, on démontre par la géométrie la relation suivante:

$$\frac{\partial \alpha}{\partial x} + \frac{\partial \beta}{\partial y} + \frac{\partial \gamma}{\partial z} = -c$$

où c désigne le double de la courbure moyenne de la surface précédente au point x, y, z considéré. Les rayons principaux doivent être comptés positivement vers l'intérieur.

La relation (6') s'écrit donc sous la forme élégante

$$\frac{dg}{dn} - cg = 2\omega^2 - 4\pi i \varrho. \tag{8}$$

Cette formule a déjà été donnée par Heinrich Bruns dans son livre «*Figur der Erde*» en 1878. Elle n'a été reproduite ni dans le traité d'Helmert, ni dans celui de Tisserand, ni même dans l'encyclopédie. M. Pizetti[1]) est le seul auteur qui la mentionne à notre connaissance. Cette formule n'est pas assez connue et nous allons voir, dans la suite, quel parti on peut en tirer. Les surfaces à ϱ constant sont à Φ constant et la formule de Bruns montre qu'une mesure de l'accroissement de g suivant la verticale équivaut à une mesure de la courbure moyenne de la surface d'égale densité passant au point où l'on opère.

Je l'ai étendue à des mouvements plus généraux que l'équilibre relatif[2]) et M. Dive[3]) l'a généralisée au cas d'une rotation permanente quelconque.

---

[1]) P. Pizetti, Principii della teoria delle figure dei pianeti, Pisa 1913.

[2]) Sur la rotation permanente des planètes, Archives des Sciences physiques et naturelles, Genève 1928.

[3]) Comptes rendus des Séances de la Société de physique de Genève 45, No. 11 928.

2. Les figures planétaires.

Nous appellerons *stratification* la répartition au point de vue strictement géométrique des surfaces d'égale densité et *densité transformée* l'expression

$$f = 2\,\omega^2 - 4\,\pi\, i\, \varrho. \tag{9}$$

Les surfaces d'égale densité seront caractérisées par un paramètre t. Les quantités ϱ, f, Φ ne dépendent que de t. Quoique les lignes de forces du champ de la pesanteur forment une multiplicité dépendant de deux paramètres, pour simplifier nos formules, nous les représenterons par un seul θ. La valeur $\theta = 0$ représentera une ligne de force particulière. Un d représentera une différentielle relative à un déplacement sur une ligne de force θ et enfin une parenthèse affectée de l'indice θ ou des indices θ', θ'', ... signifiera que la quantité entre parenthèses doit être prise le long des lignes $\theta, \theta', \theta''$, ...

Posons: $\frac{dn}{dt} = N(t, \theta)$ et puisque g dérive du potentiel Φ on aura sur une même surface t

$$(gN)_{\theta'} = (gN)_{\theta''}.$$

Dérivons par rapport à t les deux membres de cette équation et tenons compte de la formule de Bruns, on trouvera, après une réduction simple,

$$-\frac{f(t)}{(g)_\theta} = \frac{(B)_{\theta'} - (B)_{\theta''}}{(A)_{\theta'} - (A)_{\theta''}} \tag{10}$$

où l'on a posé, L désignant le logarithme népérien,

$$A = N^2 \quad \text{et} \quad B = cN + \frac{d}{dt} LN.$$

Si θ, θ', θ'' représentent trois lignes de forces distinctes, on devra avoir

$$\frac{(B)_{\theta'} - (B)_{\theta''}}{(A)_{\theta'} - (A)_{\theta''}} = \frac{(B)_{\theta'} - (B)_{\theta}}{(A)_{\theta'} - (A)_{\theta}}. \tag{11}$$

Cette équation est d'ordre purement géométrique[4]); c'est une condition nécessaire imposée à la stratification.

En vertu d'un important théorème de M. Lichtenstein[5]), les surfaces d'égale densité doivent admettre un même plan de symétrie perpendiculaire à l'axe. Supposons, en plus, que ces surfaces soient de révolution. Soient $a(t)$ et $b(t)$ le rayon polaire et le rayon équatorial de la surface t, $c_P(t)$ et $c_E(t)$ le double de la courbure moyenne de la surface t au pôle

[4]) Voir: Commentarii Mathematici Helvetici 1, p. 3, 1929.

[5]) Math. Zeitschr. 28 (1928), p. 635.

et à l'équateur, et enfin g_P la pesanteur sur l'axe polaire. L'équation (10) peut s'écrire sous la forme

$$\frac{4\pi i\varrho(t)-2\omega^2}{g_P(t)}=\frac{c_E b'-c_P a'+\frac{a''}{a'}-\frac{b''}{b'}}{b'^2-a'^2}. \tag{12}$$

Les accents représentent ici des dérivées par rapport à t. Cette relation est rigoureuse et générale. Supposons maintenant que la stratification soit formée d'ellipsoïdes concentriques et que t représente le rayon polaire de ces ellipsoïdes.

Posons $\varepsilon=b-a$. Pour des ellipsoïdes peu aplatis les courbures sont:

$$c_E=\frac{2}{t},\qquad c_P=\frac{2}{t}\left(1-2\frac{\varepsilon}{t}\right) \tag{13}$$

et l'équation (12) s'écrit pour une telle stratification

$$\frac{4\pi i\varrho(t)-2\omega^2}{g_P(t)}=\frac{1}{t}+\frac{2}{t^2}\frac{\varepsilon}{\varepsilon'}-\frac{1}{2}\frac{\varepsilon''}{\varepsilon'}. \tag{14}$$

Si ω^2 est très petit, ε, ε', ε'' le sont aussi; on peut négliger ω^2 dans le premier membre et calculer $g_P(t)$ dans l'hypothèse d'une stratification sphérique. Si $D(t)$ désigne la densité moyenne de la matière intérieure à la couche t on trouve facilement

$$\frac{4\pi i\varrho(t)}{g_P(t)}=-\frac{3}{t}-\frac{D'}{D} \tag{15}$$

et l'équation (14) s'écrit sous la forme suivante:

$$\frac{2}{t}+\frac{D'}{D}=\frac{2}{t^2}\frac{\varepsilon}{\varepsilon'}-\frac{1}{2}\frac{\varepsilon''}{\varepsilon'}. \tag{16}$$

Enfin, en introduisant l'aplatissement $e=\frac{\varepsilon}{t}=\frac{b-a}{a}$ on trouve, après un calcul tout élémentaire, l'*équation de Clairaut* qui lie l'aplatissement des couches à la densité moyenne:

$$e''D+2e'D+\frac{2}{t}eD'+\frac{6}{t}e'D=0. \tag{17}$$

Représentons maintenant par c_ϱ le double de la courbure moyenne des surfaces à ϱ constant, et soit c_U la même quantité pour les surfaces à U constant, surfaces équipotentielles pour le champ de Newton.

Sur l'axe polaire, la formule de Bruns nous donne:

$$\frac{dg}{dt}+c_\varrho g=4\pi i\varrho-2\omega^2$$

et l'équation de Poisson s'écrit au même point

$$\frac{dg}{dt}+c_U g=4\pi i\varrho.$$

En les soustrayant membre à membre, on trouve:

$$(18)\qquad \frac{2\omega^2}{g} = c_U - c_\varrho .$$

Puisque les surfaces sont de révolution, on a plus simplement:

$$(19)\qquad \frac{\omega^2}{g} = \frac{1}{R_U} - \frac{1}{R_\varrho}$$

les R étant les rayons de courbure des méridiennes. Les surfaces étant de nouveau supposées ellipsoïdales et de faible aplatissement, c_U variera entre deux limites qui correspondent: 1° au cas d'une concentration de toute la masse au centre et 2° au cas d'une répartition homogène.

On trouve pour ces deux circonstances

$$1°:\quad \frac{1}{R_U} = \frac{1}{t}, \qquad 2°:\quad \frac{1}{R_U} = \frac{1}{t} - \frac{6}{5}\frac{\varepsilon}{t}$$

tandis que l'on a toujours

$$\frac{1}{R_\varrho} = \frac{1}{t} - 2\frac{\varepsilon}{t},$$

on déduit alors de l'équation (19) les inégalités

$$\frac{4}{5}\varepsilon \leqq \frac{\omega^2 t}{g} \leqq 2\varepsilon, \qquad \frac{1}{2}\frac{\omega^2 t}{g} \leqq \varepsilon \leqq \frac{4}{5}\frac{\omega^2 t}{g}$$

qui constituent elles aussi un important *théorème de Clairaut.*

A l'extérieur de l'astre $\varrho = 0$ et la formule (14) donne en négligeant les termes du second ordre

$$4\varepsilon + 2t\varepsilon' - t^2\varepsilon'' = 0 .$$

Cette équation linéaire admet la solution

$$(20)\qquad \varepsilon = \omega^2\lambda(t^4 + \varkappa t^{-1})$$

λ et $\varkappa$ étant deux constantes. Nous allons voir que λ est lié à la masse totale de la planète et $\varkappa$ à ses moments d'inertie.

Soit S une surface de niveau extérieure à l'astre. En intégrant les deux membres de la formule (6′) à l'intérieur de S, on trouve, après usage d'une identité de Green, l'équation de Poincaré

$$(21)\qquad \iint g\,dS = 4\pi i M - 2\omega^2 T$$

où M est la masse totale de l'astre et T le volume total intérieur à S. Désignons alors par φ la latitude d'un point de S où la pesanteur est g et par g_0 la pesanteur au pôle de S. On a

$$g = \frac{g_0}{\frac{dn}{dt}} = \frac{g_0}{1 + \varepsilon'\cos^2\varphi} = g_0(1 - \varepsilon'\cos^2\varphi)$$

et d'autre part

$$T = \frac{4}{3}\pi t^3, \qquad dS = (t + \varepsilon \cos\theta)^2 d\Sigma,$$

$d\Sigma$ étant l'élément de la sphère unité.

En portant ces expressions dans la formule (21), on trouvera après intégration, et en remplaçant ε par sa valeur (20)

$$\frac{iM}{t^2} - \frac{2}{3}\omega^2 t = g_0\left[1 + 2\omega^2\lambda\left(\varkappa t^{-2} - \frac{2}{3}t^3\right)\right].$$

Puis mettons en évidence la partie principale de g_0

$$g_0 = \frac{iM}{t^2} + \omega^2 g_{0,1} \tag{22}$$

dans l'équation précédente, ce qui donne

$$-g_{0,1} = 2iM\lambda\varkappa t^{-4} + \frac{2}{3}t(1 - 2iM\lambda).$$

Le coefficient de t doit être nul car g ne peut pas croître avec t; on a donc

$$\lambda = \frac{1}{2iM} \tag{23}$$

et

$$g_{0,1} = -\varkappa t^{-4}. \tag{24}$$

Les deux relations déduites de (20)

$$\frac{\varepsilon}{t} = \omega^2\lambda(t^3 + \varkappa t^{-2}), \tag{25}$$

$$\varepsilon' = \omega^2\lambda(4t^3 - \varkappa t^{-2}) \tag{26}$$

donnent par addition

$$\frac{\varepsilon}{t} + \varepsilon' = \frac{5}{2}\frac{\omega^2 t^3}{iM} \tag{27}$$

ou

$$\frac{\varepsilon}{t} + \varepsilon' = \frac{5}{2}\frac{\omega^2 t}{g}. \tag{28}$$

C'est une relation de Clairaut qui s'écrit encore, en introduisant la pesanteur g_1 sur l'équateur de S, et l'aplatissement lui-même e,

$$e + \frac{g_1 - g_0}{g} = \frac{5}{2}\frac{\omega^2 t}{g}. \tag{29}$$

Soient A, B, C les moments d'inertie de la masse par rapport aux axes x, y, z. On aura:

$$C = \iiint \varrho(x^2 + y^2)\,dT$$

les autres expressions se déduisent de celle-là par permutation des lettres A, B, C; x, y, z. La densité ϱ satisfait à la relation

$$4\pi i\varrho = 2\omega^2 - \nabla\Phi.$$

Cette valeur introduite dans C donne après usage de l'identité de Green

$$(30)\quad 4\pi i C = 4\iiint(\Phi_S - \Phi)\,dT + \iint g(x^2+y^2)\,dS + 2\omega^2\iiint(x^2+y^2)\,dT.$$

La soustraction de deux formules semblables donne:

$$(31)\qquad 4\pi i(C-A) = \iint g(x^2-z^2)\,dS + 2\,\omega^2\iiint(x^2-z^2)\,dT.$$

Cette relation montre que $C-A$ ne dépend que de la surface S, de la vitesse angulaire et des valeurs de g sur S. Elle convient quels que soient l'aplatissement et la vitesse angulaire. Elle s'écrit encore, en désignant par T' le volume compris entre S et une sphère centrée à l'origine et de même pôle que S,

$$4\pi i(C-A) = \iint g(x^2-z^2)\,dS + 2\,\omega^2\iiint(x^2-z^2)\,dT'.$$

En négligeant les termes du second ordre comme nous l'avons fait précédemment, on trouve simplement

$$4\pi i(C-A) = \iint g(x^2-z^2)\,dS.$$

Enfin, en remplaçant g et dS par leurs valeurs on trouvera, facilement,

$$(32)\qquad C-A = \frac{t^4}{8i}\left(\frac{M}{t^2} - g_0\right),$$

ou

$$(33)\qquad C-A = 3i\omega^2\varkappa.$$

C'est la relation cherchée liant $\varkappa$ aux moments d'inertie. Enfin, en tirant $\varkappa$ de la formule (20) on trouve la *relation de d'Alembert*:

$$(34)\qquad C-A = \frac{2}{3}\,\frac{g}{i}\,t^4\left(e - \frac{\omega^2 t}{g}\right).$$

Les relations (29) ... (34) sont vraies à l'extérieur de l'astre.

On voit avec quelle rapidité la formule de Bruns permet d'obtenir les résultats classiques de la théorie des figures planétaires.

(Eingegangen am 22. 4. 1929.)

Bemerkungen zu meiner Arbeit „Über die Summen zufälliger Größen“[1].

Von

A. Kolmogoroff in Moskau.

Den Zweck dieser Bemerkungen bilden einerseits einige Berichtigungen und Präzisierungen (§§ 1 und 2), andererseits die Untersuchung eines wichtigen Spezialfalls des Gesetzes der großen Zahlen (§ 3).

§ 1.

Herrn V. W. Niemytzki verdanke ich den Hinweis, daß im Beweis des Satzes IV der oben erwähnten Abhandlung ich die falsche Formel

$$\mathfrak{J}_{\mathfrak{t}_c}(t) \geqq \frac{1}{2} D$$

benutzte. In der Tat ist leicht zu beweisen, daß

$$\mathfrak{J}_{\mathfrak{t}_c}(t) = \frac{1}{2}\,\mathfrak{D}\,|t| \leqq \frac{1}{2}\sqrt{\mathfrak{D}(t^2)} = \frac{1}{2} D$$

gilt, also die genau inverse Ungleichung.

Nun will ich die Sätze IV, V und VI durch folgende ersetzen:

Satz IV*. *Es sei $M \leqq D$. Dann ist*

$$\mathfrak{W}\left(|s| \geqq \frac{D}{2}\right) \geqq \frac{1}{1600}.$$

Satz V*.

$$\mathfrak{W}(|s| \geqq R) \geqq \frac{1}{1600}\left[1 - \frac{M^2 + 4R^2}{D^2}\right].$$

Die Sätze VII bis XI bleiben ungeändert, da wir den Satz VI nur bei dem Beweis des Satzes VIII benutzten, wo man ihn durch den neuen Satz V* ersetzen kann.

---

[1]) Math. Annalen **99** (1928), S. 309.

Beweis des Satzes IV*. Wir bezeichnen mit $\mathfrak{e}$ den Fall

$$|s| \geqq \frac{D}{2}$$

und mit $\bar{\mathfrak{e}}$ den entgegengesetzten Fall. Wenn

$$|\mathfrak{D}(s)| \geqq 2D$$

ist, so haben wir

$$D^2 \geqq \mathfrak{J}_{\bar{\mathfrak{e}}}(t^2) \geqq \mathfrak{J}_{\bar{\mathfrak{e}}}\left(2D - \frac{D}{2}\right)^2 \geqq \frac{9}{4}\mathfrak{W}(\bar{\mathfrak{e}})D^2,$$

$$\mathfrak{W}(\bar{\mathfrak{e}}) \leqq \frac{1}{2},$$

womit unser Satz bewiesen ist. Wenn

$$|\mathfrak{D}(s)| < 2D$$

ist, setzen wir

$$\mathfrak{e}_m = \left[3mD \leqq |t| < 3(m+1)D,\ |s| \geqq \frac{D}{2}\right].$$

Wegen des Satzes II haben wir

$$\mathfrak{W}(\mathfrak{e}_m) \leqq \frac{1}{2^{2m}},$$

$$\mathfrak{J}_{\mathfrak{e}_m}(s^2) \leqq \frac{(m+1)^2}{2^{2m}}25D^2,$$

und da im Fall

$$\mathfrak{e}' = \sum_{m=0}^{5}\mathfrak{e}_m$$

die Ungleichungen

$$|t| \leqq 15D, \qquad |s| \leqq 20D$$

erfüllt sind, haben wir auch

$$\mathfrak{J}_{\mathfrak{e}'}(s^2) \leqq \mathfrak{W}(\mathfrak{e})\,400D^2.$$

Da für den Fall $\bar{\mathfrak{e}}$

$$\mathfrak{J}_{\bar{\mathfrak{e}}}(s^2) \leqq \frac{1}{4}D^2$$

gilt, so haben wir endlich

$$D^2 \leqq \mathfrak{J}(s^2) \leqq \left(\mathfrak{J}_{\bar{\mathfrak{e}}} + \mathfrak{J}_{\mathfrak{e}'} + \sum_{m=6}^{\infty}\mathfrak{J}_{\mathfrak{e}_m}\right)(s^2)$$

$$\leqq \frac{1}{4}D^2 + 400D^2\mathfrak{W}(\mathfrak{e}) + \sum_{m=6}^{\infty}\frac{(m+1)^2}{2^{2m}}25D^2 \leqq \frac{3}{4}D^2 + 400D^2\mathfrak{W}(\mathfrak{e}),$$

$$\mathfrak{W}(\mathfrak{e}) \geqq \frac{1}{1600},$$

w. z. b. w.

Beweis des Satzes V*. Wenn $M > D$ oder $R > \frac{1}{2}D$ gilt, ist unser Satz richtig, da in diesen Fällen die rechte Seite der Ungleichung negativ ist. Wenn aber $M \leqq D$ und $R \leqq \frac{1}{2}D$ gilt, so erhalten wir leicht unseren Satz aus dem Satz IV*.

§ 2.

Bei dem Beweis des Satzes VIII meiner oben zitierten Abhandlung benutzte ich die Formel

$$(16) \qquad \sum_{k=1}^{m_n} \mathfrak{W}(|y_{nk} - f_{nk}| \geqq m_n n) \to 0.$$

Für den strengen Beweis dieser Formel ist folgender Satz nötig:

Satz VI*. *Es sei* $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_n$ *eine Folge unabhängiger Ereignisse und* $\mathfrak{u}$ *ein willkürliches Ereignis, für welche*

$$\mathfrak{W}_{\mathfrak{e}_k}(\mathfrak{u}) \geqq u, \qquad \mathfrak{W}(\mathfrak{e}_1 + \mathfrak{e}_2 + \ldots + \mathfrak{e}_n) \geqq u$$

gilt. Dann ist

$$\mathfrak{W}(\mathfrak{u}) \geqq \frac{1}{9} u^2.$$

Beweis des Satzes VI*. Wenn es ein solches k gibt, daß

$$\mathfrak{W}(\mathfrak{e}_k) \geqq \frac{1}{3} u$$

gilt, so haben wir

$$\mathfrak{W}(\mathfrak{u}) \geqq \mathfrak{W}(\mathfrak{e}_k)\, \mathfrak{W}\mathfrak{e}_k(\mathfrak{u}) \geqq \frac{1}{3} u^2,$$

womit unser Satz bewiesen ist.

Wenn aber für alle k

$$\mathfrak{W}(\mathfrak{e}_k) < \frac{1}{3} u$$

gilt, so gibt es gewiß ein solches k, daß

$$\frac{1}{3} u \leqq \mathfrak{W}(\mathfrak{e}_1 + \mathfrak{e}_2 + \ldots + \mathfrak{e}_k) \leqq \frac{2}{3} u$$

ist. Bezeichnen wir nun mit $\mathfrak{f}_i$ den Fall

$$\mathfrak{f}_i = \mathfrak{e}_1 + \mathfrak{e}_2 + \ldots + \mathfrak{e}_{i-1}$$

und mit $\bar{\mathfrak{f}}_i$ den entgegengesetzten Fall. Da die Ereignisse $\mathfrak{f}_i$ und $\mathfrak{e}_i$ gegenseitig unabhängig sind, erhalten wir für $i \leqq k$

$$\mathfrak{W}_{\mathfrak{e}_i}(\mathfrak{f}_i) = \mathfrak{W}(\mathfrak{f}_i) \leqq \mathfrak{W}(\mathfrak{f}_k) \leqq \frac{2}{3} u,$$

$$\mathfrak{W}_{\mathfrak{e}_i}(\mathfrak{u}\bar{\mathfrak{f}}_i) \geqq \frac{1}{3} u,$$

$$\mathfrak{W}(\mathfrak{u}\bar{\mathfrak{f}}_i\mathfrak{e}_i) \geqq \frac{1}{3} u\, \mathfrak{W}(\mathfrak{e}_i),$$

$$\mathfrak{W}(\mathfrak{u}) \geqq \sum_{i=1}^{k} \mathfrak{W}(\mathfrak{u}\bar{\mathfrak{f}}_i\mathfrak{e}_i) \geqq \frac{1}{3} u \sum_{i=1}^{k} \mathfrak{W}(\mathfrak{e}_i) \geqq \frac{1}{3} u\, \mathfrak{W}(\mathfrak{f}_k) \geqq \frac{1}{9} u^2,$$

w. z. b. w.

Beweis der Formel (16). Es sei

$$\mathfrak{u} = \left(|\sigma_n - d_n| \geqq \frac{\eta}{2}\right),$$

$$\mathfrak{e}_k = (|y_{nk} - f_{nk}| \geqq m_n \eta),$$

$$\mathfrak{f}_k = \left\{\left|\frac{\sum\limits_{i \neq k} y_{ni} + f_{nk}}{m_n} - d_n\right| \geqq \frac{\eta}{2}\right\}.$$

Es ist offenbar

$$\mathfrak{W}_{\mathfrak{f}_k}(\mathfrak{u}) \geqq \frac{1}{2}.$$

Für genügend große n haben wir, unter der Voraussetzung der Stabilität der Mittelwerte,

$$\mathfrak{W}(\mathfrak{u}) \leqq \frac{1}{4}$$

und folglich

$$\mathfrak{W}(\mathfrak{f}_k) \leqq \frac{1}{2}.$$

Wenn $\mathfrak{e}_k$ erfüllt und $\mathfrak{f}_k$ nicht erfüllt ist, so ist $\mathfrak{u}$ erfüllt. Folglich haben wir

$$\mathfrak{W}_{\mathfrak{e}_k}(\mathfrak{f}_k) + \mathfrak{W}_{\mathfrak{e}_k}(\mathfrak{u}) \geqq 1.$$

Aber da $\mathfrak{e}_k$ und $\mathfrak{f}_k$ gegenseitig unabhängig sind, so gilt für genügend große n

$$\mathfrak{W}_{\mathfrak{e}_k}(\mathfrak{f}_k) = \mathfrak{W}(\mathfrak{f}_k) \leqq \frac{1}{2},$$

$$\mathfrak{W}_{\mathfrak{e}_k}(\mathfrak{u}) \geqq \frac{1}{2}.$$

Wegen des Satzes VI* erhalten wir

$$\mathfrak{W}(\mathfrak{u}) \geqq \frac{1}{36}\mathfrak{W}^2(\mathfrak{e}_1 + \mathfrak{e}_2 + \dots + \mathfrak{e}_{m_n}),$$

und da $\mathfrak{W}(\mathfrak{u})$ gegen Null konvergiert, so gilt auch

$$\mathfrak{W}(\mathfrak{e}_1 + \mathfrak{e}_2 + \dots + \mathfrak{e}_{m_n}) \to 0,$$

$$\sum_{k=1}^{m_n} \mathfrak{W}(\mathfrak{e}_k) \to 0,$$

w. z. b. w.

§ 3.

Wir wollen nun einen besonderen Fall des Gesetzes der großen Zahlen untersuchen, nämlich den, wo alle unabhängigen Größen

$$y_1, y_2, \dots, y_n, \dots$$

dieselbe Verteilung

$$\mathfrak{W}(y < x) = F(x)$$

haben. In diesem Fall gilt:

Satz XII. *Für die Stabilität der Mittelwerte unabhängiger Größen mit einer festen Verteilung ist die Bedingung*

$$n\mathfrak{W}(|y| > n) \to 0$$

notwendig und hinreichend.

Zum Beweis, daß diese Bedingung hinreichend ist, setzen wir

$$\bar{y}_{nk} = y_n, \quad \text{wenn} \quad |y_n| \leqq n,$$
$$\bar{y}_{nk} = 0, \quad \text{wenn} \quad |y_n| > n.$$

Es ist leicht einzusehen, daß, wenn unsere Bedingung erfüllt ist, das System $\{\bar{y}_{nk}\}$ mit $\{y_n\}$ äquivalent ist, und daß

$$\frac{1}{n^2}\sum_{k=1}^{n}\mathfrak{D}(\bar{z}_{nk}^2) \leqq \frac{1}{n^2}\sum_{k=1}^{n}\mathfrak{D}(\bar{y}_{nk}^2) \to 0$$

gilt. Folglich sind wegen des Satzes VIII die Mittelwerte stabil.

Sind andererseits die Mittelwerte stabil, so gilt, wie wir bewiesen haben, die Formel (16). In unserem Falle verwandelt sich (16) in

$$n\mathfrak{W}(|y - f| > n\eta) \to 0,$$

wo f eine Konstante ist. Aus dieser Formel folgt die zu beweisende Bedingung unmittelbar.

Wenn unsere Größen y_n eine bestimmte mathematische Erwartung

$$\int_{-\infty}^{+\infty} x\, dF(x)$$

haben, so ist die Bedingung XII erfüllt, da in diesem Fall

$$\int_{-\infty}^{+\infty} |x|\, dF(x) < \infty,$$

$$n\mathfrak{W}(|y| > n) \leqq \left(\int_{-\infty}^{-n} + \int_{+n}^{-\infty}\right)|x|\, dF(x) \to 0$$

gilt. Man könnte zeigen, daß die Stabilität in diesem Falle normal sein wird. So haben wir

Satz XIII. *Für die normale Stabilität der Mittelwerte unabhängiger Größen mit fester Verteilung ist die Existenz der mathematischen Erwartung notwendig und hinreichend*[2]).

Der letzte Satz ist schon früher von A. Khintchine bewiesen worden[3]).

---

[2]) Die Notwendigkeit ist trivial und hängt damit zusammen, daß die Definition der normalen Stabilität nur im Falle der Existenz der mathematischen Erwartung einen Sinn hat.

[3]) Comptes Rendus 188, S. 477.

(Eingegangen am 8. 2. 1929.)

Über gewisse Differential- und allgemeinere Gleichungen, deren Lösungen fastperiodisch sind.

I. Teil. Der Existenzsatz.

Von

S. Bochner in München.

Inhaltsverzeichnis.

§ 1.

Formulierung der Sätze.

Die Herren H. Bohr und O. Neugebauer[1]) haben die Differentialgleichung $(-\infty < x < +\infty)$

$$(1)\qquad y^{(r)}(x) + a_1 y^{(r-1)}(x) + \dots + a_r y(x) = f(x)$$

mit konstanten Koeffizienten a_ϱ und fastperiodischer rechter Seite $f(x)$ [2]) daraufhin untersucht, inwiefern sie Lösungen besitzt, die wiederum fastperiodisch sind. In Anschluß hieran hat Herr A. Walther[3]) dieselbe Frage

[1]) Über lineare Differentialgleichungen mit konstanten Koeffizienten und fastperiodischer rechter Seite, Göttinger Nachrichten (1926), S. 8—49.

[2]) Unter einer fastperiodischen Funktion verstehen wir eine im allgemeinen komplexwertige Funktion, welche im ursprünglichen Bohrschen Sinne (also gleichmäßig stetig und) fastperiodisch ist.

[3]) Über lineare Differenzengleichungen mit konstanten Koeffizienten und fastperiodischer rechter Seite, Göttinger Nachrichten (1927), S. 196—216.

für die Differenzengleichung

$$(2)\qquad y(x+s\omega)+a_1 y(x+(s-1)\omega) + a_2 y(x+(s-2)\omega)+\dots+a_s y(x)=f(x)$$

erörtert.

Das Ziel der vorliegenden Note ist es, die Ergebnisse aus diesen Arbeiten in einheitlicher Fassung neu darzustellen und zu vertiefen.

Um die Tragweite unserer Behandlungsweise hervortreten zu lassen, werden wir unserer Diskussion eine recht allgemeine Differenzen-Differentialgleichung zugrunde legen, welche die Gleichungen (1) und (2) als Spezialfälle enthält, aber ihrerseits durch die Ansätze in den genannten Arbeiten nicht mehr erfaßt werden könnte. Und was die Resultate anbetrifft, so bemerken wir, daß in den zitierten Arbeiten im wesentlichen nur der Beschränktheitssatz (vgl. weiter unten) gewonnen wird, und daß unser Existenzsatz auch für die Spezialfälle (1) und (2) über die vorliegenden Teilergebnisse erheblich hinausgeht.

1. Die von uns betrachtete Gleichung lautet

$$(3)\qquad L(y)=f(x),$$

wobei

$$(3_1)\qquad L(y)\equiv\sum_{\varrho=0}^{r}\sum_{\sigma=0}^{s}a_{\varrho\sigma}y^{(\varrho)}(x+\delta_\sigma).$$

Hierin bezeichnen: r und s irgendwelche nicht-negative ganze Zahlen, die $a_{\varrho\sigma}$ irgendwelche (auch teilweise verschwindende) komplexe Zahlen, die δ_σ irgendwelche reelle Zahlen (und nicht, wie in (2), nur solche, welche ganzzahlige Multipla eines gemeinsamen Faktors ω sind), und $f(x)$ irgendeine fastperiodische Funktion

$$(4)\qquad f(x)\sim\sum a_{\lambda_n}e^{i\lambda_n x}.$$

Sonderfälle der allgemeinen Gleichung, die im folgenden eine Rolle spielen werden, sind: a) der Fall $r=0$, also

$$(5)\qquad L(y)\equiv\sum_{\sigma=0}^{s}a_{0\sigma}y(x+\delta_\sigma),$$

und b) für $r\geqq 1$ der Fall des dominierenden Differentialquotienten

$$(6)\qquad L(y)\equiv y^{(r)}(x)+\sum_{\varrho=0}^{r-1}\sum_{\sigma=0}^{s}a_{\varrho\sigma}y^{(\varrho)}(x+\delta_\sigma).$$

Wir bemerken, daß im Falle b) die Voraussetzung des weiter unten formulierten Existenzsatzes immer erfüllt ist.

2. Beschränktheitssatz. *Jede Lösung der Gleichung* (3), *welche mitsamt ihren r ersten Ableitungen beschränkt und gleichmäßig stetig ist, ist mitsamt ihren r ersten Ableitungen fastperiodisch.*

Eine Diskussion und der Beweis dieses Satzes werden den Inhalt des zweiten Teiles dieser Arbeit bilden.

3. Existenzsatz. *Von der linken Seite der Gleichung* (3) *sei vorausgesetzt, daß es eine Zahl* $\eta > 0$ *gibt, derart, daß für alle* λ *aus* $-\infty < \lambda < +\infty$

$$\left|\sum_{\sigma=0}^{s} a_{r\sigma} e^{i\delta_\sigma \lambda}\right| \geqq \eta. \tag{7}$$

Demzufolge (vgl. § 4, 5) *hat die Funktion*

$$G(\lambda) = \sum_{\varrho=0}^{r} \sum_{\sigma=0}^{s} a_{\varrho\sigma} (i\lambda)^{\varrho} e^{i\delta_\sigma \lambda} \tag{8}$$

höchstens endlich viele reelle Nullstellen

$$\Lambda_1, \Lambda_2, \ldots, \Lambda_k, \tag{9}$$

deren Vielfachheiten wir mit

$$l_1, l_2, \ldots, l_k$$

bezeichnen.

a) *Damit bei einer gegebenen Funktion* (4) *eine Lösung unserer Gleichung vorhanden ist, welche mitsamt ihren r ersten Ableitungen fastperiodisch ist, ist notwendig und hinreichend, daß es für jedes* ν *aus* $1 \leqq \nu \leqq k$ *eine fastperiodische Funktion gibt, deren* l_ν*-te Ableitung mit* $e^{-i\Lambda_\nu x} f(x)$ *übereinstimmt. Eine solche Lösung ist im wesentlichen eindeutig. Sie hat, sofern sie existiert, die Fourierreihe*

$$y(x) \sim \sum G^{-1}(\lambda_n) a_{\lambda_n} e^{i\lambda_n x} + \sum_{\nu=1}^{k} C_\nu e^{i\Lambda_\nu x}, \tag{10}$$

wobei die C_ν *willkürliche Konstanten sind.*

b) *Hat insbesondere, unter der Voraussetzung* (7), $G(\lambda)$ *überhaupt keine reelle Nullstelle, so gibt es für jedes* $f(x)$ *genau eine Lösung, welche mitsamt ihren r ersten Ableitungen fastperiodisch ist, nämlich*:

$$y(x) \sim \sum G^{-1}(\lambda_n)\, a_{\lambda_n} e^{i\lambda_n x}. \tag{11}$$

§ 2.

Vorbereitendes.

1. Gegeben sei eine Funktion $\Gamma(\lambda)$, welche für jeden reellen Wert der Variablen λ einen (komplexen) endlichen oder unendlichen Wert annimmt. Eine vorliegende Fourierreihe (4) sei so beschaffen[4]), daß die λ_n

[4]) Unter einer Fourierreihe verstehen wir eine trigonometrische Reihe $\sum a_{\lambda_n} e^{i\lambda_n x}$, welche die Fourierreihe einer fastperiodischen Funktion ist.

keine Unendlichkeitsstellen von $\Gamma(\lambda)$ sind, und daß die Reihe

$$\sum \Gamma(\lambda_n) a_{\lambda_n} e^{i\lambda_n x} \tag{12}$$

wiederum eine Fourierreihe ist. Dann werden wir $\Gamma(\lambda)$ einen ***Multiplikator*** der Reihe (4) bzw. der entsprechenden Funktion $f(x)$ nennen. Die zu (12) gehörige Funktion werden wir $f_\Gamma(x)$ schreiben.

2. Die Summe von endlich vielen Multiplikatoren von $f(x)$ ist wiederum ein Multiplikator von $f(x)$.

3. Zwei Funktionen $\Gamma_1(\lambda)$ und $\Gamma_2(\lambda)$, welche für die Exponenten λ_n von (4) gleiche Werte haben, sind entweder alle beide Multiplikatoren von (4) oder keine von beiden. Wir nennen sie „äquivalent" in bezug auf (4).

4. Es gibt Funktionen $\Gamma(\lambda)$, die in bezug auf jede fastperiodische Funktion Multiplikatoren sind. Wir nennen sie Multiplikatoren (schlechthin). Von dieser Art ist $\gamma(\lambda) = e^{i\delta\lambda}$, ($\delta$ reell): $f_\gamma(x) \equiv f(x+\delta)$. Das Produkt $\Gamma_1(\lambda)\Gamma_2(\lambda)$ zweier Multiplikatoren $\Gamma_1(\lambda)$ und $\Gamma_2(\lambda)$ ist, wie man sofort feststellt, wiederum ein Multiplikator.

5. Einen Multiplikator bildet jede Funktion $\Gamma(\lambda)$, welche stetig differenzierbar ist, und außerhalb eines genügend großen Intervalles verschwindet,

$$\Gamma(\lambda) = 0 \quad \text{für} \quad |\lambda| \geqq \alpha. \tag{13}$$

Bevor wir dies kurz beweisen, ziehen wir hieraus die Folgerung, daß jede Funktion $\Gamma^*(\lambda)$, welche in einem abgeschlossenen Intervall

$$\alpha \leqq \lambda \leqq \beta \tag{14}$$

stetig differenzierbar ist, für jede (fastperiodische) Funktion, deren Exponenten in (14) liegen, einen Multiplikator darstellt. Denn man kann $\Gamma^*(\lambda)$ außerhalb von (14) zu einer stetig differenzierbaren Funktion $\Gamma(\lambda)$ ergänzen, welche außerhalb eines passenden (größeren) Intervalls verschwindet. In bezug auf Funktionen, deren Exponenten in (14) liegen, sind aber $\Gamma(\lambda)$ und $\Gamma^*(\lambda)$ äquivalent (vgl. 3.).

Nun zum Beweis. Wir setzen mit der Hilfsfunktion

$$K(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ix\lambda}\,\Gamma(\lambda)\,d\lambda \tag{15}$$

den Ausdruck an

$$F(x) = \int_{-\infty}^{+\infty} f(t)K(t-x)\,dt = \int_{-\infty}^{+\infty} f(t+x)K(t)\,dt. \tag{16}$$

Die Existenz des uneigentlichen Integrals (16), durch welches $F(x)$ gegeben ist, wird dadurch gewährleistet, daß für die durch (15) definierte Funktion $K(x)$

das Integral

$$\int_{-\infty}^{+\infty} |K(x)|\,dx \tag{17}$$

einen endlichen Wert hat[5]). Von der Funktion $F(x)$ ist unschwer ein-

---

[5]) Dies beruht auf dem folgenden wichtigen Satz von Plancherel über Fouriersche Transformierte (wegen der einfachsten Beweise vgl. E. C. Titchmarsh, Hankel Transforms, Proc. Cambridge Philos. Soc. 21 (1923), S. 463—473 und F. Riesz, Sur la formule d'inversion de Fourier, Acta Szeged 3 (1927), S. 235—241). Es sei irgendeine Funktion $\varphi(\lambda)$ gegeben, welche auf jedem endlichen Intervall nach Lebesgue meßbar ist, und für welche darüber hinaus das Lebesguesche Integral

$$\int_{-\infty}^{+\infty} |\varphi(\lambda)|^2\,d\lambda \tag{18}$$

einen endlichen Wert hat. Dann wird durch das Integral

$$H(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ix\lambda}\,\varphi(\lambda)\,d\lambda \tag{19}$$

eine Funktion von x erzeugt, welche ihrerseits gleichfalls meßbar ist und ein endliches Integral

$$\int_{-\infty}^{+\infty} |H(x)|^2\,dx \tag{20}$$

besitzt. Es besteht die zu (18) reziproke Beziehung

$$\varphi(\lambda) = \int_{-\infty}^{+\infty} e^{i\lambda t}\,H(t)\,dt. \tag{21}$$

Die Integrale (19) bzw. (21) brauchen nicht im gewöhnlichen Lebesgueschen Sinne als uneigentliche Integrale zu existieren. Sie sind aber gemäß einer allgemeineren, den vorliegenden Verhältnissen angepaßten Definition vorhanden, durch welche jedoch die resultierenden Funktionen $H(x)$ bzw. $\varphi(\lambda)$ nur abgesehen von einer Punktmenge vom Lebesgueschen Maße Null festgelegt sind. Wenn aber etwa in (19) die vorgegebene Funktion $\varphi(\lambda)$ stetig ist und die dann resultierende Funktion $H(x)$ im Intervall $-\infty < x < +\infty$ absolut integrierbar ist, so ist die reziproke Beziehung (21) im *gewöhnlichen* Sinne und außerdem für *alle* Punkte λ gültig.

Es liege nunmehr eine differenzierbare Funktion $\Gamma(\lambda)$ vor, welche mitsamt ihrer Ableitung $\Gamma'(\lambda)$ ein endliches Integral (18) besitzt. Die Funktionen (15) und

$$K_1(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ix\lambda}\,\Gamma'(\lambda)\,d\lambda \tag{22}$$

stehen dann, auf Grund einer zulässigen partiellen Integration von (22), in der Beziehung

$$K(x) = \frac{1}{ix} K_1(x).$$

(Fortsetzung der Fußnote [5]) auf nächster Seite.)

zusehen, daß sie fastperiodisch ist, und daß

(23) $$F(x) = f_\Gamma(x).$$ [6])

6. Wegen (13) hat $f_\Gamma(x)$, wie auch $f(x)$ beschaffen sein mag, beschränkte Exponenten, und hat demnach nach einem allgemeinen Satz unbegrenzt viele fastperiodische Ableitungen. Das erkennt man auch an der Darstellung (16). Denn $K(x)$ ist beliebig oft differenzierbar,

$$K^{(\varrho)}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ix\lambda}(-i\lambda)^{\varrho}\,\Gamma(\lambda)\,d\lambda \qquad (\varrho = 1, 2, 3, \ldots),$$

und $K^{(\varrho)}(x)$ besitzt, aus demselben Grunde wie für $K(x)$ das Integral (17) einen endlichen Wert hat, ein endliches Integral

$$\int_{-\infty}^{+\infty} |K^{(\varrho)}(x)|\,dx.$$

Nach der Schwarzschen Ungleichung ist

$$\left|\int_1^{\infty} K(x)\,dx\right|^2 \leq \int_1^{\infty}\frac{1}{x^2}\,dx \int_1^{\infty} |K_1(x)|^2\,dx$$

und analog

$$\left|\int_{-\infty}^{-1} K(x)\,dx\right|^2 \leq \int_{-\infty}^{-1}\frac{1}{x^2}\,dx \int_{-\infty}^{-1} |K_1(x)|^2\,dx,$$

und auf Grund dessen existiert das Integral (17).

[6]) Auf Grund von (21) ist für irgendein λ

$$\Gamma(\lambda)\,a_\lambda e^{i\lambda x} = \int_{-\infty}^{+\infty}\{a_\lambda e^{i\lambda(t+x)}\}\,K(t)\,dt;$$

also besteht die Relation (23) für eine „Einzelschwingung" $a_\lambda e^{i\lambda x}$, und hieraus folgt sofort, daß sie für irgendein Exponentialpolynom $p(x)$ besteht. Eine allgemeine fastperiodische Funktion $f(x)$ kann man gleichmäßig durch eine Folge von Exponentialpolynomen

$$p_1(x),\ p_2(x),\ \ldots,\ p_n(x),\ \ldots$$

approximieren. Wie aus der Beziehung

$$\begin{aligned} |f_\Gamma(x) - p_\Gamma(x)| &\leq \int |f(x+t) - p(x+t)|\cdot|K(t)|\,dt \\ &\leq \underset{-\infty<\xi<+\infty}{\text{Ob. Gr.}} |f(\xi) - p(\xi)|\cdot\int |K(t)|\,dt \end{aligned}$$

zu ersehen ist, ist dann die Folge $p_{n\Gamma}(x)$ gleichmäßig gegen $f_\Gamma(x)$ konvergent. Da die Fourierreihe von $p_{n\Gamma}(x)$ aus der von $p_n(x)$ durch Multiplikation mit $\Gamma(\lambda)$ hervorgeht, und da, wegen der gleichmäßigen Konvergenz, die Fourierreihen von $p_n(x)$ bzw. $p_{n\Gamma}(x)$ mit wachsendem n Glied für Glied in die Fourierreihen von $f(x)$ bzw. $f_\Gamma(x)$ übergehen, entsteht auch die Fourierreihe von $f_\Gamma(x)$ aus der Fourierreihe von $f(x)$ durch Multiplikation mit $\Gamma(\lambda)$.

Hieraus schließt man, daß das erste Integral in (16) beliebig oft nach x differenziert werden kann, und daß die Relation besteht

$$f_\Gamma^{(\varrho)}(x) = (-1)^\varrho \int_{-\infty}^{+\infty} f(t)\, K^{(\varrho)}(t-x)\, dt. \tag{24}$$

Weiterhin gilt[7])

$$\lim K(x) = 0 \quad \text{für} \quad |x| \to \infty,$$

und analog

$$\lim K^{(\varrho)}(x) = 0 \quad \text{für} \quad |x| \to \infty \qquad (\varrho = 1, 2, 3, \ldots).$$

Wenn nunmehr $f(x)$ r fastperiodische (und daher beschränkte) Ableitungen besitzt, kann man (24), für $\varrho = 1, 2, \ldots, r$, schrittweise durch partielle Integration umformen zu

$$f_\Gamma^{(\varrho)}(x) = \int_{-\infty}^{+\infty} f^{(\varrho)}(t)\, K(t-x)\, dt = \int_{-\infty}^{+\infty} f^{(\varrho)}(t+x)\, K(t)\, dt;$$

oder anders geschrieben

$$\frac{d^{(\varrho)}}{dx^\varrho}\{f_\Gamma(x)\} = \left\{\frac{d^{(\varrho)} f(x)}{dx^\varrho}\right\}_\Gamma \qquad (\varrho = 1, 2, \ldots, r).$$

7. Allgemeiner als in 5. ist eine Funktion $\Gamma(\lambda)$ bereits dann ein Multiplikator, wenn sie etwa eine stetige Ableitung hat, und wenn in bezug auf das Verhalten im Unendlichen nur vorausgesetzt wird, daß $\int_{-\infty}^{+\infty} |\Gamma(\lambda)|^2 d\lambda$ und $\int_{-\infty}^{+\infty} |\Gamma'(\lambda)|^2 d\lambda$ endlich sind[8]), was z. B. immer dann erfüllt ist, wenn für $|\lambda| \to \infty$

$$\begin{aligned} \Gamma(\lambda) &= O(|\lambda|^{-1}) \\ \Gamma'(\lambda) &= O(|\lambda|^{-1}). \end{aligned} \tag{25}$$

8. Gegeben seien vier Konstanten $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, für welche

$$\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4. \tag{26}$$

Man kann ohne weiteres eine stetig differenzierbare Funktion $\Gamma(\lambda)$ konstruieren, welche außerhalb des Intervalls $\alpha_1 \leqq \lambda \leqq \alpha_4$ verschwindet, und in $\alpha_2 \leqq \lambda \leqq \alpha_3$ den Wert 1 hat. Nach 5. ist eine solche Funktion ein Multiplikator.

9. Von einer fastperiodischen Funktion (4) setzen wir voraus, daß sie für passende Konstante α_ν, für welche (26) erfüllt ist, zwei Lücken (α_1, α_2) und (α_3, α_4) aufweist, d. h. daß keiner ihrer Exponenten in die Intervalle

---

[7]) Dies ist, vgl. (13) und (15), der bekannte Riemannsche Satz über die Abnahme der Fourierkoeffizienten einer beliebigen Funktion gegen Null.

[8]) Vgl. [5]) und [6]).

$\alpha_1 \leqq \lambda \leqq \alpha_2$ und $\alpha_3 \leqq \lambda \leqq \alpha_4$ hineinfällt. Dann ist die Teilreihe

$$\sum_{\alpha_1 \leqq \lambda_n \leqq \alpha_2} a_{\lambda_n} e^{i\lambda_n x}$$

der Fourierreihe (4) wiederum eine Fourierreihe. Denn sie entsteht aus (4) durch Multiplikation mit der Funktion $\Gamma(\lambda)$ aus 8.

§ 3.

Die einfachste Gleichung.

Dies ist die Differentialgleichung

$$y^{(r)}(x) = f(x),$$

d. h.

$$y(x) = \underbrace{\int^x \cdots \int^x}_{r} f(\xi)\, d\xi;$$

sie bezieht sich auf mehrfache Differentiation bzw. mehrfache Integration fastperiodischer Funktionen. Hierzu wollen wir in diesem Paragraphen einige Bemerkungen zusammenstellen.

1. Wenn die Exponenten einer fastperiodischen Funktion beschränkt sind, so besitzt die Funktion unendlich viele Ableitungen, und sie sind alle fastperiodisch.

2. Ist die Ableitung von (4) vorhanden und fastperiodisch, so entsteht ihre Fourierreihe durch Multiplikation mit $i\lambda$,

$$f'(x) \sim \sum i\lambda_n a_{\lambda_n} e^{i\lambda_n x}.$$

3. Ist das Integral von (4) fastperiodisch, so entsteht es durch „Division" mit $i\lambda$,

$$\int^x f(\xi)\, d\xi \sim C_0 + \sum (i\lambda_n)^{-1} a_{\lambda_n} e^{i\lambda_n x}.$$

Wir nennen es „normiert", wenn die willkürliche additive Konstante C_0 den Wert Null hat,

$$\int^x f(\xi)\, d\xi \sim \sum (i\lambda_n)^{-1} a_{\lambda_n}^{*} e^{i\lambda_n x}.$$

Damit die Division mit $i\lambda$ ausführbar sein soll, muß in erster Linie a_0 verschwinden. Aber das reicht nicht aus. Es können noch die „kleinen" λ_n stören. Wenn es aber ein $\eta > 0$ gibt, so daß $|\lambda_n| \geqq \eta$ für alle n, so ist das Integral fastperiodisch[9]). Induktiv findet man: Falls $|\lambda_n| \geqq \eta > 0$, so

[9]) Vgl. des Verfassers: Properties of Fourier Series of almost periodic Fonctions, Proc. London Math. Soc. (2) 26 (1927), S 433—452, insbesondere S. 442ff.; oder J. Favard, Sur les fonctions harmoniques presque périodiques, Journal de Mathématiques (9) 6 (1927), S. 229—336, insbesondere S. 309ff.

ist das beliebig oft iterierte (bei jedem Schritt normierte) Integral von (4) wieder fastperiodisch,

$$\underbrace{\int^x \cdots \int^x}_{\varrho} f(\xi)\, d\xi \sim \sum (i\lambda_n)^{-\varrho} a_{\lambda_n} e^{i\lambda_n x}.$$

4. Hieraus folgt: Wenn es für zwei Fourierreihen

$$f(x) \sim \sum a_{\lambda_n} e^{i\lambda_n x},$$
$$g(x) \sim \sum b_{\lambda_n} e^{i\lambda_n x}$$

ein Intervall $\Lambda - \alpha \leqq \lambda \leqq \Lambda + \beta$ (Λ beliebig, $\alpha > 0, \beta > 0$) gibt, so daß für alle λ_n aus diesem Intervall $a_{\lambda_n} = b_{\lambda_n}$, so ist die Reihe

$$\sum \{i(\lambda_n - \Lambda)\}^{-\varrho} (a_{\lambda_n} - b_{\lambda_n}) e^{i\lambda_n x} \tag{27}$$

für positives ganzes ϱ eine Fourierreihe. Denn sie entsteht aus

$$h(x) \sim \sum (a_{\lambda_n} - b_{\lambda_n}) e^{i\lambda_n x}$$

durch die sukzessiven Bildungen

$$e^{-i\Lambda x} h(x) \sim \sum (a_{\lambda_n} - b_{\lambda_n}) e^{i(\lambda_n - \Lambda)x},$$

$$h_\varrho(x) = \underbrace{\int^x \cdots \int^x}_{\varrho} e^{-i\Lambda\xi} h(\xi)\, d\xi \sim \sum \{i(\lambda_n - \Lambda)\}^{-\varrho} (a_{\lambda_n} - b_{\lambda_n}) e^{i(\lambda_n - \Lambda)x},$$

$$e^{i\Lambda x} h_\varrho(x) \sim (27),$$

von denen nach 3. auch die mittlere die Fastperiodizität besitzt.

5. Die Reihe (4) habe für passendes ganzzahliges $r > 0$ den Divisor $(i\lambda)^r$,

$$F(x) \sim \sum (i\lambda_n)^{-r} a_{\lambda_n} e^{i\lambda_n x}.$$

Dann sind das 1-te, 2-te, ..., r-te Integral von $f(x)$, in passender Normierung, wieder fastperiodisch. Um das einzusehen, betrachte man irgendein $\Gamma(\lambda)$ gemäß § 2, 8., mit $\alpha_2 < 0 < \alpha_3$, bilde die Funktionen $f_1(x) = f_\Gamma(x)$ und $F_1(x) = F_\Gamma(x)$, und setze

$$f(x) = f_1(x) + f_2(x),$$
$$F(x) = F_1(x) + F_2(x).$$

Nach 1. ist $F_1(x)$ beliebig oft differenzierbar, nach 2. ist

$$F_1^{(r)}(x) = f_1(x),$$

also hat $f_1(x)$ r fastperiodische Integrale; nach 3. besitzt $f_2(x)$ unbegrenzt viele fastperiodische Integrale; also besitzt $f(x)$ fastperiodische Integrale bis zur r-ten Ordnung.

6. Wenn man die Rollen von $f(x)$ und $F(x)$ vertauscht und 2. berücksichtigt, erhält man folgendes. Damit die Funktion (4) den Multipli-

kator $(i\lambda)^r$ besitzt, ist notwendig und hinreichend, daß sie fastperiodische Ableitungen bis zur r-ten Ordnung hat.

Aus einem Satz von Esclangon folgt, daß eine beschränkte Funktion $g(x)$ $(-\infty < x < +\infty)$, welche ein beschränktes r-tes Integral besitzt, auch ein beschränktes 1-tes, 2-tes, ..., $(r-1)$-tes Integral besitzt. In Verbindung mit dem Satz, daß ein beschränktes Integral einer fastperiodischen Funktion selber fastperiodisch ist, schließen wir hieraus (mit Bohr und Neugebauer loc. cit. S. 14), daß eine fastperiodische Funktion, welche ein fastperiodisches r-tes Integral hat, auch ein fastperiodisches 1-tes, 2-tes, ..., $(r-1)$-tes Integral besitzt.

7. Nunmehr können wir folgende Behauptung aussprechen. Es sei Λ eine reelle Zahl und r eine natürliche Zahl. Damit zugleich mit der Reihe (4) auch die Reihe

$$\sum \{i(\lambda_n - \Lambda)\}^{-r} a_{\lambda_n} e^{i\lambda_n x}$$

eine Fourierreihe ist, ist notwendig und hinreichend, daß die Funktion

$$e^{-i\Lambda x} f(x)$$

ein fastperiodisches r-tes Integral besitzt.

§ 4.

Die homogene Gleichung.

1. Eine (unbestimmte) fastperiodische Lösung $y(x)$ von (3) werden wir in der Gestalt

$$y(x) \sim \sum A_{\lambda_n} e^{i\lambda_n x} \tag{28}$$

ansetzen und voraussetzen, daß in (4) und (28) die Fourierexponenten die gleichen sind, weil man zwei ganz beliebige Fourierreihen durch Einschalten verschwindender Fourierterme auf gleiche Exponenten bringen kann. Falls die r ersten Ableitungen von (28) fastperiodisch sind, so ist für beliebiges reelles δ_ϱ und für $1 \leqq \varrho \leqq r$

$$y^{(\varrho)}(x + \delta_\varrho) \sim \sum (i\lambda_n)^\varrho e^{i\lambda_n \delta_\varrho} A_{\lambda_n} e^{i\lambda_n x},$$

und daher

$$L(y) \sim \sum G(\lambda_n) A_{\lambda_n} e^{i\lambda_n x}, \tag{29}$$

wo $G(\lambda)$ durch (8) definiert ist. Wenn (3) eine mit den r ersten Ableitungen fastperiodische Lösung (28) besitzen soll, so muß für alle n die Relation

$$G(\lambda_n) A_{\lambda_n} = a_{\lambda_n} \tag{30}$$

bestehen, d. h. es muß $G(\lambda)$ ein Divisor von (4) sein. Daher spielen die (reellen) Nullstellen von $G(\lambda)$ eine wichtige Rolle. Da $G(\lambda)$ eine ganze Funktion der komplexen Veränderlichen λ ist, hat sie höchstens abzählbar viele reelle Nullstellen

$$(31) \qquad \Lambda_1, \Lambda_2, \Lambda_3, \ldots,$$

und es ist $\lim\limits_{k \to \infty} |\Lambda_k| = \infty$. Die Vielfachheiten der Nullstellen seien

$$(31_1) \qquad l_1, l_2, l_3, \ldots .$$

Damit die Gleichung (30) erfüllt sein kann, darf in (4) kein Exponent wesentlich (d. h. mit einem nichtverschwindenden Koeffizienten) vorkommen, der einer Zahl aus (31) gleicht.

2. Sind *alle* $a_{\lambda_n} = 0$, ist also die Gleichung (3) homogen, so sind genau alle diejenigen mitsamt den ersten r Ableitungen fastperiodischen Funktionen Lösungen, deren Exponenten der Folge (31) entnommen sind. Ist die Folge (31) leer, so gibt es dann überhaupt keine solche Lösung.

3. In einer Lösung (28) der inhomogenen Gleichung (3) können Exponenten vorkommen, die Werte aus der Folge (31) annehmen. Wenn es eine Lösung gibt, in der *kein* Exponent Λ_ν wesentlich vorkommt, wollen wir sie die *reine* Lösung nennen. Gibt es überhaupt eine Lösung (28), und ist die Folge (31) endlich, so gibt es die reine Lösung. Man erhält sie, wenn man von der Funktion (28) das Exponentialpolynom

$$(32) \qquad \sum A_{\Lambda_\nu} e^{i\Lambda_\nu}$$

abzieht.

4. Es sei dahingestellt, ob dies auch im Falle einer unendlichen Folge (31) zutrifft, d. h. ob die mit den Fourierkoeffizienten A_λ einer Lösung (28) gebildete Reihe (32) wiederum eine Fourierreihe ist. Für die Gleichung (2) ist dies nach Walther loc. cit. zutreffend. Es trifft auch für die allgemeine Gleichung (3) immer dann zu, wenn die wesentlichen Exponenten von (28) unterhalb einer festen Schranke liegen, $|\lambda_n| \leqq \eta$. Weil nämlich die Λ_ν isoliert sind, kann man nach § 2, 9. von der Funktion $y(x)$ die Reihe

$$\sum_{|\lambda_n| \gtrless \eta} A_{\lambda_n} e^{i\lambda_n x}$$

als Fourierreihe abtrennen; und diese Fourierreihe repräsentiert eine Funktion, die fastperiodische Ableitungen beliebig hoher Ordnung besitzt, vgl. § 3, 1. Also ist die Reihe

$$\sum_{|\lambda_n| \geqq \eta} A_{\lambda_n} e^{i\lambda_n x} = \sum_{|\Lambda_\nu| \geqq \eta} A_{\Lambda_\nu} e^{i\Lambda_\nu x}$$

und demnach auch die Gesamtreihe (32) eine fastperiodische Funktion, welche zugleich mit (28) fastperiodische Ableitungen bis zur r-ten Ordnung aufweist.

5. Im Falle (7) ist die Folge (31) endlich, weil dann, wegen $|e^{i\delta\lambda}|=1$, die Funktion $G(\lambda)$ außerhalb eines angebbaren beschränkten Intervalls von Null verschieden ist.

§ 5.

Beweis des Existenzsatzes.

1. Im vorliegenden Paragraphen werden wir nur den Absatz b) des Existenzsatzes beweisen. Der Beweis ad a) wird sich dann im Verlauf des nächsten Paragraphen ergeben.

2. Wir behandeln zuerst den Spezialfall (5), also

$$G(\lambda)=g(\lambda)=\sum_{\sigma=0}^{s} a_{0\sigma}e^{i\delta_\sigma\lambda}. \tag{33}$$

Nach (30) kommt, wegen $G(\lambda)\neq 0$, als Lösung nur eine einzige fastperiodische Funktion in Frage, nämlich eine eventuelle Funktion mit der Fourierreihe

$$\sum g^{-1}(\lambda_n)\,a_{\lambda_n}e^{i\lambda_n x}.$$

Und wie man leicht einsieht, läuft unser Satz auf die Behauptung hinaus, daß

$$\gamma(\lambda)=g^{-1}(\lambda)$$

ein Multiplikator (§ 2, 4.) ist. Da $g(\lambda)$ ein Exponentialpolynom, also eine fastperiodische Funktion ist, so ist wegen (7) auch $\gamma(\lambda)$ als Funktion der reellen Variablen λ eine fastperiodische Funktion,

$$\gamma(\lambda)\sim\sum\gamma_\nu e^{i\mu_\nu\lambda}. \tag{34}$$

Diese Funktion besitzt nun eine für unsern jetzigen Zusammenhang wichtige Eigenschaft: ihre Fourierreihe ist *absolut* konvergent[10]). Jede Partialsumme

$$\gamma(N)=\gamma(N;\lambda)=\sum_{\nu=0}^{N}\gamma_\nu e^{i\mu_\nu\lambda}$$

ist nach § 2, 4. und § 2, 2. ein Multiplikator:

$$f_{\gamma(N)}(x)=\sum_{\nu=0}^{N}\gamma_\nu\sum_n a_{\lambda_n}e^{i\lambda_n(\mu_\nu+x)}=\sum_{\nu=0}^{N}\gamma_\nu f(x+\mu_\nu).$$

Wegen der Konvergenz von $\sum\limits_{\nu=0}^{\infty}|\gamma_\nu|$ ist der in x gleichmäßige Limes

$$\lim_{N=\infty}\sum_{\nu=0}^{N}\gamma_\nu f(x+\mu_\nu)$$

---

[10]) Wegen des Beweises vgl. des Verfassers: Beitrag zur absoluten Konvergenz fastperiodischer Fourierreihen, Jahresbericht d. Deutsch. Mathematiker-Vereinigung (im Druck).

vorhanden, ist als solcher für eine fastperiodische Funktion $f(x)$ wiederum fastperiodisch, und seine Fourierreihe beträgt (nach der Regel über die formale Konvergenz von Fourierreihen)

$$\sum_n \lim_{N\to\infty} \gamma(N; \lambda_n) a_{\lambda_n} e^{i\lambda_n x} = \sum_n \gamma(\lambda_n) a_{\lambda_n} e^{i\lambda_n x}.$$

Also ist, was zu beweisen war, $\gamma(\lambda)$ ein Multiplikator der willkürlich vorgegebenen fastperiodischen Funktion $f(x)$.

3. Nunmehr behandeln wir den Spezialfall (6), also

$$G(\lambda) = g_*(\lambda) = (i\lambda)^r + \sum_{\varrho=0}^{r-1} \sum_{\sigma=0}^{s} (i\lambda)^\varrho e^{i\delta_\sigma \lambda}. \tag{35}$$

Von den Funktionen

$$(i\lambda)^\varrho g_*^{-1}(\lambda) \qquad (\varrho = 0, 1, \ldots, r-1) \tag{36}$$

verifiziert man leicht, daß sie den Bedingungen (25) genügen, also nach § 2, 7. Multiplikatoren sind. Nach § 2, 4. und § 2, 2. ist auch

$$1 - g_*^{-1}(\lambda) \sum_{\varrho=0}^{r-1} \sum_{\sigma=0}^{s} (i\lambda)^\varrho e^{i\delta_\sigma \lambda} = (i\lambda)^r g_*^{-1}(\lambda)$$

ein Multiplikator. Es existiert daher die fastperiodische Funktion

$$\sum g_*^{-1}(\lambda_n) a_{\lambda_n} e^{i\lambda_n x},$$

und wegen § 3, 6. sind die r ersten Ableitungen vorhanden und fastperiodisch. Es ist dies die im Satz behauptete Lösung.

4. Für die allgemeinste Funktion $G(\lambda)$, welche den Voraussetzungen im Absatz b) des Existenzsatzes genügt, sind wiederum die Funktionen

$$(i\lambda)^\varrho G^{-1}(\lambda) \qquad (\varrho = 0, 1, \ldots, r-1),$$

weil sie den Bedingungen (25) genügen, Multiplikatoren; also ist es auch die Funktion

$$1 - G^{-1}(\lambda) \sum_{\varrho=0}^{r-1} \sum_{\sigma=0}^{s} (i\lambda)^\varrho e^{i\delta_\sigma \lambda}. \tag{36_1}$$

Sie hat den Wert

$$g(\lambda) \cdot (i\lambda)^r G^{-1}(\lambda), \tag{36_2}$$

wo $g(\lambda)$ wie in (33), mit Koeffizienten $a_{r\sigma}$ statt $a_{0\sigma}$, definiert ist. Wir benutzen nunmehr das Resultat aus 2., daß $g^{-1}(\lambda)$ ein Multiplikator ist, und daß daher, vgl. § 2, 4., auch

$$(i\lambda)^r G^{-1}(\lambda)$$

von dieser Beschaffenheit ist. Jetzt schließt man wie in 3.

§ 6.

Ein weiterer Satz.

1. Satz. *Es liege die allgemeinste Gleichung* (3) *vor und die dazugehörige Funktion* $G(\lambda)$ *habe die* (*endlich oder unendlich vielen*) *Nullstellen* (31) *mit den Vielfachheiten* (31_1).

a) *Damit es eine Lösung* (28) *von* (3) *geben kann, welche mitsamt ihren r ersten Ableitungen fastperiodisch ist, ist notwendig, daß für jedes* ν *eine fastperiodische Funktion existiert, deren* l_ν*-te Ableitung mit* $e^{-i\Lambda_\nu x} f(x)$ *übereinstimmt.*

b) *Falls die wesentlichen Fourierexponenten von* $f(x)$ *beschränkt sind, ist diese Bedingung auch hinreichend, und die Lösung lautet*

$$y(x) \sim \sum G^{-1}(\lambda_n)\, a_{\lambda_n} e^{i\lambda_n x} + \sum A_{\Lambda_\nu} e^{i\Lambda_\nu x}, \tag{37}$$

wobei der zweite Summand eine beliebige mitsamt den r ersten Ableitungen fastperiodische Funktion bedeutet, deren Exponenten der Folge (31) *entnommen sind.*

2. Beweis ad a). Jeder Nullstelle Λ_ν ordnen wir Zahlen $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ mit der Bedingung (26) zu, derart, daß Λ_ν ins Intervall $\alpha_2 \leqq \lambda \leqq \alpha_3$ fällt und das Intervall $\alpha_1 \leqq \lambda \leqq \alpha_4$ keine von Λ_ν verschiedene Zahl aus (31) enthält; und wir konstruieren irgendeinen Multiplikator, wir nennen ihn $\Gamma_\nu(\lambda)$, gemäß § 2, 8. mit den vorliegenden Zahlen α.

Die zwischen der Funktion $y(x)$ und $f(x)$ bestehende Relation (3) überträgt sich nunmehr, auf Grund von § 2, 6., für jedes ν auf die Funktionen

$$y_\nu(x) \sim \sum \Gamma_\nu(\lambda_n) A_{\lambda_n} e^{i\lambda_n x} \tag{38}$$

und

$$f_\nu(x) \sim \sum \Gamma_\nu(\lambda_n) a_{\lambda_n} e^{i\lambda_n x}, \tag{39}$$

nämlich

$$L(y_\nu) = f_\nu(x). \tag{40}$$

Nach Einführung der Abkürzungen

$$A_\lambda^{(\nu)} = \Gamma_\nu(\lambda) A_\lambda \quad \text{und} \quad a_\lambda^{(\nu)} = \Gamma_\nu(\lambda) a_\lambda \tag{41}$$

lautet (40) folgendermaßen:

$$\sum G(\lambda_n) A_{\lambda_n}^{(\nu)} e^{i\lambda_n x} = \sum a_{\lambda_n}^{(\nu)} e^{i\lambda_n x}. \tag{41_1}$$

Die Funktion

$$G_\nu(\lambda) = \{i(\lambda - \Lambda_\nu)\}^{-l_\nu} \cdot G(\lambda) \tag{42}$$

ist überall stetig differenzierbar, daher ist sie, da die Exponenten von $y_\nu(x)$ beschränkt sind, auf Grund von § 2, 5. ein Multiplikator von $y_\nu(x)$. Da-

her existiert die Fourierreihe

$$\sum G_\nu(\lambda_n) A_{\lambda_n}^{(\nu)} e^{i\lambda_n x}.$$

Die Relation (41_1) besagt, daß das Produkt dieser Fourierreihe mit

$$\{i(\lambda - \Lambda_\nu)\}^{l_\nu} \tag{43}$$

die Fourierreihe von $f_\nu(x)$ ergibt, daß also (43) ein Divisor von $f_\nu(x)$ ist. Zugleich mit $f_\nu(x)$ besitzt aber, auf Grund von § 3, 4., auch $f(x)$ selbst den Divisor (43). Der Beweis von a) ergibt sich nunmehr aus § 3, 7.

3. Es sei jetzt vorausgesetzt, daß für ein festes ν die Funktion $e^{-i\Lambda_\nu x} f(x)$ ein l_ν-tes fastperiodisches Integral besitzt. Durch Kombination von § 3, 7. und § 3, 4. erhalten wir, daß die Reihe

$$\sum_n \{i(\lambda_n - \Lambda_\nu)\}^{-l_\nu} a_{\lambda_\nu}^{(\nu)} e^{i\lambda_n x} \tag{44}$$

eine Fourierreihe ist. Ihre wesentlichen Exponenten liegen in einem abgeschlossenen Intervall, in welchem $G_\nu(\lambda)$ von Null verschieden ist; daher kann man, vgl. § 2, 5., die Reihe (44) mit $G_\nu(\lambda)$ dividieren. Die resultierende Reihe lautet, wenn wir $A_\lambda = G^{-1}(\lambda) a_\lambda$ setzen,

$$y_\nu(x) \sim \sum A_{\lambda_n}^{(\nu)} e^{i\lambda_n x} \tag{45}$$

und die so herausspringende Funktion $y_\nu(x)$ ist, wie man sofort verifiziert, eine (mitsamt allen Ableitungen) fastperiodische Lösung von (40).

4. Beweis ad b). Die wesentlichen Exponenten von $f(x)$ mögen im Intervall

$$-\eta \leqq \lambda \leqq +\eta \tag{46}$$

gelegen sein. Wir betrachten diejenigen Zahlen aus (31), welche in das Intervall (46) fallen, und nehmen an, daß es etwa die Zahlen $\Lambda_1, \Lambda_2, \ldots, \Lambda_k$ sind. Wir setzen nunmehr

$$\Gamma_0(\lambda) = 1 - \Gamma_1(\lambda) - \Gamma_2(\lambda) - \ldots - \Gamma_k(\lambda) \tag{47}$$

und ziehen die Definitionen (38), (39) und (41) auch für den Wert $\nu = 0$ heran. Die Exponenten von $f_0(x)$ sind wegen (46) beschränkt und zwar auf endlich viele beschränkte abgeschlossene Intervalle verteilt, innerhalb welcher $G(\lambda)$ von Null verschieden ist. Also existiert die Funktion

$$y_0(x) \sim \sum G^{-1}(\lambda_n) a_{\lambda_n}^{(0)} e^{i\lambda_n x};$$

sie befriedigt die Gleichung (40) für $\nu = 0$. Da nun

$$f(x) = f_0(x) + f_1(x) + \ldots + f_k(x)$$

ist, so ist

$$y(x) = \sum_0^k y_\nu(x) \sim \sum G^{-1}(\lambda_n) a_{\lambda_n} e^{i\lambda_n x}$$

eine Funktion von beschränkten Exponenten, welche der Gleichung (3) genügt. Daß die allgemeine Lösung die Gestalt (37) hat, folgt nunmehr nach § 4, 2.

5. **Beweis zum Absatz a) des Existenzsatzes.**

Daß die Bedingung notwendig ist, ist bereits in 3. bewiesen worden. Es sei nunmehr umgekehrt bekannt, daß für jede Zahl aus (9) die Funktion $e^{-i\Lambda_\nu x} f(x)$ ein fastperiodisches l_ν-tes Integral besitzt. Wir bilden wiederum wie in 2. und gemäß (47) die Funktionen $\Gamma_\nu(\lambda)$, $0 \leqq \nu \leqq k$, und betrachten wiederum die Gleichungen (40). Für $\nu \neq 0$ erhalten wir wiederum die Lösungen (45). Für $\nu = 0$ berücksichtigen wir, daß die wesentlichen Exponenten von $f_0(x)$ in endlich viele abgeschlossene Intervalle fallen (von denen zwei unendlich sind), so daß in diesen Intervallen $G(\lambda)$ von Null verschieden ist. In den komplementären offenen Intervallen, welche an Zahl endlich und jedes beschränkt sind, denken wir uns $G(\lambda)$ irgendwie zu einer stetig differenzierbaren Funktion $G_0(\lambda)$ abgeändert, welche nirgends verschwindet. Die Funktion $G_0^{-1}(\lambda)$ ist in bezug auf $f_0(x)$ der ursprünglichen Funktion $G^{-1}(\lambda)$ äquivalent (§ 2, 3.). Auf diese Funktion übertragen sich aber sehr leicht[11]) die Überlegungen aus § 5, 4. und ergeben eine mitsamt den r ersten Ableitungen fastperiodische Lösung

$$y_0(x) \sim \sum G_0^{-1}(\lambda_n) a_{\lambda_n} e^{i\lambda_n x}$$

der Gleichung (40) für $\nu = 0$. Jetzt schließt man zu Ende, wie in 4.

---

[11]) Man muß jetzt $g(\lambda)$ durch die Gleichung

$$(36_1) = (36_2)$$

definieren.

(Eingegangen am 16. 3. 1929.)

Über das monotone Polynom, welches die minimale Abweichung von Null hat, wenn die Werte seiner ersten Ableitungen gegeben sind.

Von

W. Břečka und J. Geronimus in Charkow.

Wir wollen das folgende Problem behandeln:

Man soll in dem Intervall $(-1, +1)$ *die minimale Abweichung von Null des Polynoms höchstens n-ten Grades*

$$y_n(x) = A_0 x^n + A_1 x^{n-1} + \ldots + A_n$$

finden, welches in diesem Intervall monoton wachsend ist, wenn die Werte seiner k ersten Ableitungen in einem Punkte $|\eta| < 1 - \varepsilon$ *gegeben sind*:

$$y_n^{(i)}(\eta) = a_i \qquad (i = 1, 2, \ldots, k),$$

wobei ε *eine beliebig kleine, aber von n unabhängige Zahl bedeutet.*

I. Kapitel.

§ 1.

Wir wollen mit einem leichteren Problem beginnen:

Man soll das Minimum des Integrals

$$S = \int_{-1}^{1} q(x)\, R_m^2(x)\, dx$$

finden, wo $q(x)$ *ein gegebenes Polynom l-ten Grades ist, welches keine Wurzel in dem Intervall* $-1 < x < 1$ *hat*; $R_m(x)$ *ist ein Polynom m-ten Grades. Die Werte dieses Polynoms und seiner r ersten Ableitungen in einem Punkte* $|\eta| < 1 - \varepsilon$ *sind gegeben*:

$$R_m^{(i)}(\eta) = b_i \qquad (i = 0, 1, 2, \ldots, r).$$

Um dieses Problem zu lösen, wollen wir

$$R_m(x) = \sum_{s=0}^{m} c_s P_s(x)$$

setzen, wo die Polynome $P_s(x)$ normiert und orthogonal sind, d. h.

$$\int_{-1}^{1} q(x) P_s(x) P_p(x)\, dx = \begin{cases} 0 & \text{für } s \neq p, \\ 1 & \text{für } s = p. \end{cases}$$

Es soll nun

$$S = \sum_{s=0}^{m} c_s^2 \tag{1}$$

unter den Nebenbedingungen

$$\sum_{s=0}^{m} c_s P_s^{(i)}(\eta) = b_i \qquad (i = 0, 1, 2, \ldots, r)$$

zum Minimum gemacht werden.

Dann erhalten wir die Gleichungen

$$2c_s = \sum_{i=0}^{r} \lambda_i P_s^{(i)}(\eta) \qquad (s = 0, 1, 2, \ldots, m), \tag{2}$$

und hieraus folgt

$$2S = \sum_{i=0}^{r} \lambda_i b_i. \tag{3}$$

Weiter multiplizieren wir die Gleichungen (2) mit $P_s^{(j)}(\eta)$, summieren und erhalten

$$2b_j = \sum_{i=0}^{r} \lambda_i a_{ij} \qquad (j = 0, 1, 2, \ldots, r), \tag{4}$$

wo

$$a_{ij} = \sum_{s=0}^{m} P_s^{(i)}(\eta) P_s^{(j)}(\eta)$$

ist.

Nach (3) und (4) finden wir

$$\begin{vmatrix} S & b_0 & b_1 & \ldots & b_r \\ b_0 & a_{00} & a_{10} & \ldots & a_{r0} \\ b_1 & a_{01} & a_{11} & \ldots & a_{r1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ b_r & a_{0r} & a_{1r} & \ldots & a_{rr} \end{vmatrix} = 0$$

und endlich

$$S = \sum_{i=0}^{r} \sum_{j=0}^{r} (-1)^{i+j} b_i b_j \frac{A_{ij}}{a}, \tag{4'}$$

wo

$$a = \begin{vmatrix} a_{00} & a_{10} & \dots & a_{r0} \\ a_{01} & a_{11} & \dots & a_{r1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{0r} & a_{1r} & \dots & a_{rr} \end{vmatrix}$$

ist und A_{ij} die Unterdeterminante dieser Determinante ist, welche dem Element a_{ij} entspricht.

§ 2.

Jetzt soll

$$a_{ij} = \sum_{s=0}^{m} P_s^{(i)}(\eta)\, P_s^{(j)}(\eta) \tag{5}$$

gefunden werden. Die Formel

$$\sum_{s=0}^{m} P_s(x)\, P_s(y) = c_m \frac{P_{m+1}(x)\, P_m(y) - P_m(x)\, P_{m+1}(y)}{x-y}\ ^{1)}, \tag{6}$$

wo

$$c_m = \frac{d_m^{(m)}}{d_{m+1}^{(m+1)}}, \qquad (P_m(x) = d_m^{(m)} x^m + d_{m-1}^{(m)} x^{m-1} + \dots + d_0^{(m)}),$$

benutzend, erhalten wir nach i-maliger Differentiation von (6) in bezug auf x folgendes:

$$\sum_{s=0}^{m} P_s^{(i)}(x)\, P_s(y) = c_m \left\{ P_m(y) \sum_{r=0}^{i} (-1)^r \binom{i}{r} \frac{P_{m+1}^{(i-r)}(x)\, r!}{(x-y)^{r+1}} \right.$$
$$\left. - P_{m+1}(y) \sum_{r=0}^{i} (-1)^r \binom{i}{r} \frac{P_m^{(i-r)}(x)\, r!}{(x-y)^{r+1}} \right\},$$

und indem wir alsdann diesen Ausdruck j-mal in bezug auf y differenzieren, erhalten wir

$$\sum_{s=0}^{m} P_s^{(i)}(x)\, P_s^{(j)}(y) = c_m \left\{ \sum_{r=0}^{i} (-1)^r \binom{i}{r} P_{m+1}^{(i-r)}(x) \sum_{l=0}^{j} \binom{j}{l} \frac{P_m^{(j-l)}(y)\,(r+l)!}{(x-y)^{r+l+1}} \right.$$
$$\left. - \sum_{r=0}^{i} (-1)^r \binom{i}{r} P_m^{(i-r)}(x) \sum_{l=0}^{j} \binom{j}{l} \frac{P_{m+1}^{(j-l)}(y)\,(r+l)!}{(x-y)^{r+l+1}} \right\}.$$

Setzen wir $x = y + t$ und entwickeln alsdann alle Polynome, die von x abhängig sind, in eine Taylorreihe und setzen schließlich $t = 0$, so finden

[1]) Vgl. Darboux, Approximation des fonctions de très grands nombres, Journal de Mathématiques pures et appliquées (3) 4 (1878), p. 413.

wir, daß unsere Summe den Wert

$$\sum_{s=0}^{m} P_s^{(i)}(\eta) P_s^{(j)}(\eta)$$

$$= c_m \sum_{l=0}^{j} \left\{ \binom{j}{l} [P_m^{(j-l)}(\eta) P_{m+1}^{(i+l+1)}(\eta) - P_{m+1}^{(j-l)}(\eta) P_m^{(i+l+1)}(\eta)] \sum_{r=0}^{i} \frac{(-1)^r \binom{i}{r}}{l+r+1} \right\}$$

hat; es ist aber

$$\sum_{r=0}^{i} (-1)^r \binom{i}{r} \cdot \frac{1}{l+r+1} = \sum_{r=0}^{i} (-1)^r \binom{i}{r} \int_0^1 x^{r+l}\, dx = \int_0^1 (1-x)^i x^l\, dx$$

$$= B(i+1, l+1) = \frac{i!\, l!}{(i+l+1)!}$$

und wir erhalten endlich

$$\sum_{s=0}^{m} P_s^{(i)}(\eta) P_s^{(j)}(\eta) \tag{7}$$

$$= c_m \frac{i!\, j!}{(i+j+1)!} \sum_{l=0}^{j} \binom{i+j+1}{j-l} [P_m^{(j-l)}(\eta) P_{m+1}^{(i+l+1)}(\eta) - P_{m+1}^{(j-l)}(\eta) P_m^{(i+l+1)}(\eta)]$$

$$= c_m \frac{i!\, j!}{(i+j+1)!} \sum_{p=0}^{j} \binom{i+j+1}{p} [P_m^{(p)}(\eta) P_{m+1}^{(i+j+1-p)}(\eta) - P_{m+1}^{(p)}(\eta) P_m^{(i+j+1-p)}(\eta)].$$

§ 3.

Wir wollen die asymptotische Form von (7) finden, indem wir annehmen, daß m ins Unendliche wächst (i und j bleiben endlich).

Die Polynome $R_n(x)$, welche die Bedingungen der Orthogonalität

$$\int_{-1}^{1} \frac{p^2(x)}{\sqrt{1-x^2}} R_m(x) R_n(x)\, dx = 0 \qquad (m \neq n)$$

erfüllen, lassen die asymptotischen Ausdrücke

$$R_m(x) \sim \sqrt{\frac{2}{\pi}} \cdot \frac{\cos(m\varphi + \psi)}{p(x)}$$

zu, wo $\varphi = \arccos x$ und ψ eine Funktion ist, die von $p(x)$ abhängt[2]).
Setzt man

$$q(x) = \frac{p^2(x)}{\sqrt{1-x^2}}, \qquad \eta = \cos\varphi \qquad (|\eta| < 1 - \varepsilon),$$

---

[2]) S. Bernstein, Sur quelques propriétés asymptotiques de la meilleure approximation, Comptes Rendus **186**, p. 840. — Sur les polynomes orthogonaux, ibidem **188**, p. 361.

so erhält man

$$P_m(\cos\varphi) \sim \sqrt{\frac{2}{\pi}} \cdot \frac{\cos(m\varphi+\psi)}{\sqrt{q(\cos\varphi)\sin\varphi}}.$$

Alsdann ist

$$P_m^{(p)}(\cos\varphi) \sim \frac{\sqrt{2}\cos\left(m\varphi+\psi+p\cdot\frac{\pi}{2}\right)}{\sqrt{\pi q(\cos\varphi)\sin\varphi}\,(-\sin\varphi)^p}\cdot m^p$$

und

$$P_m^{(p)}(\cos\varphi)\,P_{m+1}^{(i+j+1-p)}(\cos\varphi) - P_{m+1}^{(p)}(\cos\varphi)\,P_m^{(i+j+1-p)}(\cos\varphi)$$

$$\sim \frac{2}{\pi}\cdot\left\{\frac{\cos\left(m\varphi+\psi+p\frac{\pi}{2}\right)\cos\left[(m+1)\varphi+\psi+(i+j+1-p)\frac{\pi}{2}\right]}{(-1)^{i+j+1}q(\cos\varphi)(\sin\varphi)^{i+j+2}}\right.$$

$$\left. - \frac{\cos\left[(m+1)\varphi+\psi+p\frac{\pi}{2}\right]\cos\left[m\varphi+\psi+(i+j+1-p)\frac{\pi}{2}\right]}{(-1)^{i+j+1}q(\cos\varphi)(\sin\varphi)^{i+j+2}}\right\}\cdot m^{i+j+1}$$

$$= \frac{m^{i+j+1}}{\pi}\left\{\frac{\cos\left[(2m+1)\varphi+2\psi+(i+j+1)\frac{\pi}{2}\right]+\cos\left[\varphi+(i+j+1)\frac{\pi}{2}-p\pi\right]}{(-1)^{i+j+1}q(\cos\varphi)(\sin\varphi)^{i+j+2}}\right.$$

$$\left. - \frac{\cos\left[(2m+1)\varphi+2\psi+(i+j+1)\frac{\pi}{2}\right]+\cos\left[\varphi-(i+j+1)\frac{\pi}{2}+p\pi\right]}{(-1)^{i+j+1}q(\cos\varphi)(\sin\varphi)^{i+j+2}}\right\}$$

$$= -\frac{2m^{i+j+1}(-1)^{i+j+1}\sin\varphi\cdot\sin\left(\frac{i+j+1}{2}\pi-p\pi\right)}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+2}}$$

$$= \frac{2m^{i+j+1}(-1)^{i+j+p}\sin\frac{i+j+1}{2}\pi}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}} = \frac{2m^{i+j+1}(-1)^{i+j}\cos\frac{i+j}{2}\pi}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}}(-1)^p.$$

Dann ist

$$a_{ij} = \sum_{s=0}^{m} P_s^{(i)}(\eta)\,P_s^{(j)}(\eta) \sim c_m\frac{i!\,j!}{(i+j+1)!}\sum_{p=0}^{j}\binom{i+j+1}{p}\frac{2m^{i+j+1}(-1)^{i+j}\cos\frac{i+j}{2}\pi\cdot(-1)^p}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}}$$

$$= c_m\frac{i!\,j!}{(i+j+1)!}\cdot\frac{2m^{i+j+1}(-1)^{i+j}\cos\frac{i+j}{2}\pi}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}}\sum_{p=0}^{j}\binom{i+j+1}{p}(-1)^p.$$

Man findet leicht, daß

$$(-1)^p\binom{i+j+1}{p} = f(p) - f(p-1)$$

ist, wo

$$f(p) = (-1)^p\binom{i+j}{p}$$

ist. Sodann ist

$$\sum_{p=0}^{j}\binom{i+j+1}{p}(-1)^p = 1+\sum_{p=1}^{j}\binom{i+j+1}{p}(-1)^p = 1+f(j)-f(0) = (-1)^j\binom{i+j}{j}$$

und

$$a_{ij} \sim \frac{2c_m(-1)^{i+j}\cos\frac{i-j}{2}\pi\cdot m^{i+j+1}}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}(i+j+1)}.$$

Die Szegösche Formel[3]) benutzend finden wir leicht, daß $c_m \sim \frac{1}{2}$ ist.

Also endlich

$$a_{ij} \sim \frac{(-1)^{i+j}\cos\frac{i-j}{2}\pi\cdot m^{i+j+1}}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}(i+j+1)}. \tag{8}$$

§ 4.

Es soll nun $\frac{A_{ij}}{a}$ gefunden werden, wo a die aus a_{ij} gebildete Determinante und A_{ij} die Unterdeterminante dieser Determinante ist, welche dem Element a_{ij} entspricht.

Man kann leicht finden, daß

$$\frac{a}{A_{ij}} \sim \frac{(-1)^{i+j}\cdot m^{i+j+1}}{\pi q(\cos\varphi)(\sin\varphi)^{i+j+1}}\cdot\frac{d}{D_{ij}}$$

ist, wo d die aus $d_{ij} = \dfrac{\cos\frac{i-j}{2}\pi}{i+j+1}$ gebildete Determinante und D_{ij} die entsprechende Unterdeterminante ist.

Nun betrachten wir das System der linearen Gleichungen:

$$\sum_{s=0}^{r} x_s \frac{\cos\frac{t-s}{2}\pi}{t+s+1} = \varepsilon_t \qquad \begin{cases}\text{für } t \neq j & \varepsilon_t = 0,\\ \text{für } t = j & \varepsilon_t = 1.\end{cases} \tag{9}$$

Dann ist

$$x_i = (-1)^{i+j}\frac{D_{ij}}{d}.$$

Es sei

$$y_i = (-1)^{\frac{i}{2}} x_i,$$

dann ist

$$\sum_{s=0}^{r}\frac{y_s}{t+s+1} = (-1)^{\frac{t}{2}}\varepsilon_t, \qquad s \equiv t \pmod 2.$$

---

[3]) G. Szegö, Über die Entwickelung einer analytischen Funktion nach den Polynomen eines Orthogonalsystems, Math. Annalen 82 (1920), S. 207.

Alsdann ist

$$y_i = (-1)^{i+j+\frac{1}{2}j} \cdot \frac{\bar{D}_{ij}}{\bar{d}},$$

wo $\bar{d}$ die aus $\bar{d}_{ij} = \dfrac{1+(-1)^{i+j}}{2(i+j+1)}$ gebildete Determinante und $\bar{D}_{ij}$ die entsprechende Unterdeterminante ist.

Dann ist

$$(-1)^{i+j+\frac{1}{2}j} \cdot \frac{\bar{D}_{ij}}{\bar{d}_{ij}} = (-1)^{i+j+\frac{1}{2}i} \cdot \frac{D_{ij}}{d}$$

und

$$\frac{D_{ij}}{d} = (-1)^{\frac{i-j}{2}} \cdot \frac{\bar{D}_{ij}}{\bar{d}}.$$

Die Werte von $\dfrac{\bar{D}_{ij}}{\bar{d}}$ sind[4])

$$\frac{\bar{D}_{ij}}{\bar{d}} = \frac{(\nu+i+1)(\nu+j+1)\bar{P}_\nu^{(i)}(0)\,\bar{P}_\nu^{(j)}(0)}{i!\; j!\; (i+j+1)} \begin{cases} \nu = r & \text{für } i \equiv j \equiv r \\ \nu = r-1 & \text{für } i \equiv j \equiv r-1 \end{cases} (\operatorname{mod} 2),$$

wo $\bar{P}_\nu(x)$ das Legendresche Polynom ist. Für $i \equiv j - 1 \pmod 2$ ist $\bar{D}_{ij} = 0$. Also haben wir

$$\frac{D_{ij}}{d} = (-1)^{\frac{i-j}{2}} \cdot \frac{(\nu+i+1)(\nu+j+1)\,\bar{P}_\nu^{(i)}(0)\,\bar{P}_\nu^{(j)}(0)}{i!\; j!\; (i+j+1)}.$$

und endlich

$$(10)\quad S \sim \pi q(\cos\varphi) \sum_{i=0}^{r} \sum_{j=0}^{r} \frac{(\sin\varphi)^{i+j+1}(-1)^{\frac{i-j}{2}}\, b_i b_j}{i!\; j!\; (i+j+1)\; m^{i+j+1}} (\nu+i+1)(\nu+j+1)\,\bar{P}_\nu^{(i)}(0)\,\bar{P}_\nu^{(j)}(0)$$

$$= \frac{\pi q(\cos\varphi)}{2^{2\nu}} \sum_{i=0}^{r} \sum_{j=0}^{r} \frac{(\sin\varphi)^{i+j+1} b_i b_j}{m^{i+j+1}} \cdot \frac{(\nu+i+1)!\; (\nu+j+1)!}{\left(\frac{\nu+i}{2}\right)!\left(\frac{\nu-i}{2}\right)!\left(\frac{\nu+j}{2}\right)!\left(\frac{\nu-j}{2}\right)!\; i!\; j!\; (i+j+1)}$$

Wenn alle Zahlen b_i von einer und derselben Ordnung sind, so ist

$$(11)\qquad S \sim \frac{\pi b_0^2\, q(\cos\varphi)\sin\varphi}{m} \cdot \frac{(\nu+1)!^2}{2^{2\nu}\left(\frac{\nu}{2}\right)!^4},$$

wo $\nu = r$, *wenn* r *gerade, und*

$\nu = r - 1$, *wenn* r *ungerade ist.*

---

[4]) J. Geronimus, Über die minimale mittlere quadratische Abweichung des Polynoms von Null im gegebenen Intervall (Russisch), Communications de la Société Mathématique de Kharkow (4) 2 (1928), S. 32.

Wenn $\frac{b_i}{i!} \sim \beta_i b_0 m^i$, *wo* β_i *endliche Zahlen sind, so ist*

$$S \sim \frac{\pi b_0^2 q(\cos\varphi)}{2^{2\nu} m} \sum_{i=0}^{r} \sum_{j=0}^{r} \frac{(\sin\varphi)^{i+j+1} \beta_i \beta_j (\nu+i+1)!\,(\nu+j+1)!}{(i+j+1)\left(\frac{\nu+i}{2}\right)!\left(\frac{\nu-i}{2}\right)!\left(\frac{\nu+j}{2}\right)!\left(\frac{\nu-j}{2}\right)!}, \tag{12}$$

wo
$$\nu = r \quad \textit{für} \quad i \equiv j \equiv r \pmod 2,$$
$$\nu = r-1 \quad \textit{für} \quad i \equiv j \equiv r-1 \pmod 2.$$

II. Kapitel.

§ 5.

Jetzt wollen wir zum monotonen Polynom zurückkehren. Es ist klar, daß *es folgende Form hat*

$$y_n(x) = \int_{-1}^{x} u^2(x) q(x) (1-x)^\alpha (1+x)^\beta \, dx \qquad (\alpha, \beta = 0, 1),$$

wo $q(x)$ *keine Wurzel für* $-1 \leqq x \leqq 1$ *hat.*

Man kann beweisen, daß $q(x)$ *höchstens* $(k-1)$*-ten Grades ist, wenn* k *gerade, und* k*-ten Grades, wenn* k *ungerade ist.*

Ist k gerade, so läßt sich ein Polynom konstruieren

$$\bar{y}(x) = \int_{-1}^{x} u^2(x)(1-x)^\alpha(1+x)^\beta [x-\eta]^k \, dx;$$

dann erfüllt, wenn $q(x)$ vom Grade $\geqq k$ ist, das Polynom

$$R_n(x) = y_n(x) - \lambda \bar{y}(x) = \int_{-1}^{x} u^2(x)(1-x)^\alpha (1+x)^\beta \; [q(x) - \lambda(x-\eta)^k] \, dx$$

alle Bedingungen unseres Problems, wenn λ so klein gewählt ist, daß

$$q(x) - \lambda(x-\eta)^k > 0 \quad \text{für} \quad -1 \leqq x \leqq 1;$$

aber für $\lambda > 0$ ist

$$R_n(1) < y_n(1).$$

Folglich ist $q(x)$ höchstens $(k-1)$-ten Grades. Ist k ungerade, so konstruiert man das Polynom auf folgende Weise

$$z(x) = \int_{-1}^{x} u^2(x)(1-x)^\alpha(1+x)^\beta [x-\eta]^{k+1} \, dx$$

und man beweist, daß $q(x)$ höchstens k-ten Grades ist.

Demnach können wir

$$y_n(x) = \int_{-1}^{x} q(x)\, u^2(x)\, dx$$

setzen, wo der Grad von $q(x)$ *endlich ist.*

Es sei

$$q(x) = (x-\alpha_1)(x-\alpha_2)\ldots(x-\alpha_k).$$

Dann haben wir

$$\frac{q'(\eta)}{q(\eta)} = \sum_{s=1}^{k} \frac{1}{\eta-\alpha_s}.$$

Wir sehen, daß $\left|\frac{q'(\eta)}{q(\eta)}\right|$ eine endliche Zahl ist oder sogar, wenn alle Wurzeln α_i unendlich groß sind, gegen Null strebt.

Ebenso haben wir

$$\frac{q''(\eta)}{q(\eta)} = \left\{\frac{q'(\eta)}{q(\eta)}\right\}^2 - \sum_{s=1}^{k} \frac{1}{(\eta-\alpha_s)^2}$$

usw. So sehen wir, daß für $s > 0$

$$\lim \left\{\frac{1}{m^s} \frac{q^{(i)}(\eta)}{q(\eta)}\right\} = 0 \qquad (i = 1, 2, 3, \ldots, k;\ m \to \infty).$$

Wenn jedoch einige Wurzeln den Wert ± 1 haben, z. B. $\alpha_i = 1$, so haben wir

$$\left|\frac{1}{\eta - a_i}\right| < \frac{1}{\varepsilon},$$

wo ε eine kleine positive Konstante ist, die von n unabhängig ist. Deshalb haben

$$\frac{q'(\eta)}{q(\eta)},\quad \frac{q''(\eta)}{q(\eta)},\quad \ldots,\quad \frac{q^{(k)}(\eta)}{q(\eta)}$$

eine obere Grenze, die von n unabhängig ist.

Hiernach läßt sich unsere Aufgabe in folgender Weise auf die in §§ 1 – 4 gelöste zurückführen. Wir wissen, daß

$$q(\eta)u^2(\eta) = a_1, \quad \frac{d^t}{dx^t}[q(x)u^2(x)]_{x=\eta} = a_{t+1} \qquad (t = 1, 2, 3, \ldots, k-1)$$

ist. *Nehmen wir an, daß alle gegebenen Ableitungen a_i einer und derselben Ordnung sind und daß $a_1 \neq 0$ ist.* Dann haben wir

$$a_{s+1} = \sum_{i=0}^{s} \binom{s}{i} q^{(s-i)}(\eta) \sum_{p=0}^{i} \binom{i}{p} u^{(p)}(\eta)\, u^{(i-p)}(\eta)$$

und

$$\frac{a_{s+1}}{a_1} = \sum_{i=0}^{s} \binom{s}{i} \frac{q^{(s-i)}(\eta)}{q(\eta)} \sum_{p=0}^{i} \left\{\binom{i}{p} \frac{u^{(p)}(\eta)}{u(\eta)} \cdot \frac{u^{(i-p)}(\eta)}{u(\eta)}\right\}. \tag{13}$$

Wenn wir

$$u^{(p)}(\eta) = b_p$$

setzen, so kann man leicht ersehen, daß wir imstande sind aus diesen Gleichungen alle b_p aufzufinden, und daß alle diese Ableitungen einer und derselben Ordnung sind.

Zum Beispiel

$$q(\eta)u^2(\eta) = q(\eta)b_0^2 = a_1,$$

$$q'(\eta)u^2(\eta) + 2q(\eta)u(\eta)u'(\eta) = b_0^2 q'(\eta) + 2b_0 b_1 q(\eta) = a_2,$$

woraus

$$\frac{a_2}{a_1} = 2\frac{b_1}{b_0} + \frac{q'(\eta)}{q(\eta)}, \qquad 2\frac{b_1}{b_0} = \frac{a_2}{a_1} - \frac{q'(\eta)}{q(\eta)}.$$

Ferner ist

$$u^2(\eta)q''(\eta) + 4u(\eta)u'(\eta)q'(\eta) + 2q(\eta)u(\eta)u''(\eta) + 2q(\eta)u'^2(\eta)$$
$$= b_0^2 q''(\eta) + 4b_0 b_1 q'(\eta) + 2b_0 b_2 q(\eta) + 2b_1^2 q(\eta) = a_3$$
$$\frac{a_3}{a_1} = \frac{q''(\eta)}{q(\eta)} + 4\frac{b_1}{b_0}\cdot\frac{q'(\eta)}{q(\eta)} + 2\frac{b_2}{b_0} + 2\left(\frac{b_1}{b_0}\right)^2,$$

woraus

$$2\frac{b_2}{b_0} = \frac{a_3}{a_1} - \frac{q''(\eta)}{q(\eta)} - 2\frac{q'(\eta)}{q(\eta)}\left(\frac{a_2}{a_1} - \frac{q'(\eta)}{q(\eta)}\right) - \frac{1}{2}\left(\frac{a_2}{a_1} - \frac{q'(\eta)}{q(\eta)}\right)^2 \quad \text{usw.}$$

Es soll nun das Integral

$$L = \int_{-1}^{1} q(x)u^2(x)\,dx$$

unter den Nebenbedingungen

$$u^{(i)}(\eta) = b_i \qquad (i = 0, 1, 2, \ldots, k-1)$$

(wenn alle b_i einer und derselben Ordnung sind) zum Minimum gemacht werden.

Die Formel (11) benutzend, erhalten wir

$$L \sim \frac{\pi q(\cos\varphi)\,b_0^2 \sin\varphi}{m}(\nu+1)^2[\overline{P}_\nu(0)]^2 = \frac{\pi q(\cos\varphi)}{m}\,b_0^2 \sin\varphi \left\{\frac{(\nu+1)!}{2^\nu\left(\frac{\nu}{2}\right)!^2}\right\}^2,$$

wo $\nu = k-1$, *wenn k ungerade ist, und* $\nu = k-2$ *für gerades* k. *Endlich im Falle, wo die gegebenen Ableitungen einer und derselben Ordnung sind, ist die minimale gesuchte Abweichung von Null durch die Formel*

$$L \sim \frac{2\pi a_1 \sin\varphi}{n}\left\{\frac{(\nu+1)!}{2^\nu\left(\frac{\nu}{2}\right)!^2}\right\}^2 \tag{14}$$

gegeben, weil

$$q(\cos\varphi)\,b_0^2 = a_1$$

ist.

Wenn $k = 1$, $\nu = 0$, so finden wir, daß

$$L \sim \frac{2\pi a_1 \sin\varphi}{n} \tag{15}$$

ist[5]).

[5]) Vgl. W. Břečka and J. Geronimus, On a monotonic polynomial having the minimal deflection from zero, Tôhoku Mathematical Journal **30** (1929), Nr. 3, 4, S. 389.

§ 6.

Nun wollen wir den Fall betrachten, wo die Ordnung der gegebenen Ableitungen in geometrischer Progression wächst.

Alsdann ist

$$a_{i+1} \sim i!\, a_1 \alpha_i m^i,$$

wo α_i endliche Zahlen sind. Wir erhalten aus (13)

$$s!\,\alpha_s m^s = \sum_{i=0}^{s} \binom{s}{i} \frac{q^{(s-i)}(\eta)}{q(\eta)} \sum_{p=0}^{i} \left\{ \binom{i}{p} \frac{b_p}{b_0} \cdot \frac{b_{i-p}}{b_0} \right\}. \tag{16}$$

Wir wollen zeigen, daß

$$\frac{b_i}{i!} \sim b_0 \beta_i m^i$$

ist. Wir sehen, daß

$$\frac{a_2}{a_1} = \frac{q'(\eta)}{q(\eta)} + 2\frac{b_1}{b_0}$$

ist, also

$$\frac{b_1}{1!} \sim b_0 \beta_1 m,$$

weil

$$\frac{a_2}{a_1} \sim \alpha_1 m$$

ist. *Wir wollen beweisen, daß, wenn*

$$\frac{b_i}{i!} \sim b_0 \beta_i m^i \qquad (i = 1, 2, 3, \ldots, s-1)$$

ist, so

$$\frac{b_s}{s!} \sim b_0 \beta_s m^s$$

ist.

Aus (16) finden wir, daß

$$s!\,\alpha_s m^s = \sum_{i=0}^{s-1} i! \binom{s}{i} \frac{q^{(s-i)}(\eta)}{q(\eta)} \left\{ \sum_{p=0}^{i} \beta_p \beta_{i-p} m^i \right\} + \sum_{p=0}^{s} \binom{s}{p} \frac{b_p}{b_0} \cdot \frac{b_{s-p}}{b_0}$$

ist. Die zweite Summe ist gleich

$$2\frac{b_s}{b_0} + s! \sum_{p=1}^{s-1} \beta_p \beta_{s-p} m^s,$$

also ist

$$\frac{b_s}{s!} \sim b_0 \beta_s m^s$$

und

$$2\beta_s = \alpha_s - \sum_{p=1}^{s-1} \beta_p \beta_{s-p}. \tag{17}$$

Diese Formel gestattet alle β zu berechnen.

Es soll nun das Integral

$$L = \int_{-1}^{1} q(x)\, u^2(x)\, dx$$

unter den Nebenbedingungen

$$u^{(i)}(\eta) = b_i \qquad (i = 0, 1, 2, \ldots, k-1),$$

wo

$$\frac{b_i}{i!} \sim b_0 \beta_i m^i$$

ist, zum Minimum gemacht werden.

Aus der Formel (12) erhalten wir

$$(18) \qquad L \sim \frac{\pi a_1}{2^{2\nu-1} n} \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \frac{(\sin\varphi)^{i+j+1} \beta_i \beta_j (\nu+i+1)!\,(\nu+j+1)!}{(i+j+1)\left(\frac{\nu+i}{2}\right)!\left(\frac{\nu-i}{2}\right)!\left(\frac{\nu+j}{2}\right)!\left(\frac{\nu-j}{2}\right)!},$$

wo $i \equiv j \pmod 2$

und $\nu = k-1$, wenn k ungerade ist,

und $\nu = k-2$ für gerades k.

Die Werte von β können aus den Gleichungen (17) berechnet werden.

Charkow (Ukraine).

(Eingegangen am 30. 3. 1929.)

Zusatz zum vorangehenden Artikel der Herren W. Břečka und J. Geronimus über monotone Polynome minimaler Abweichung.

Von

Serge Bernstein in Charkow.

Im vorangehenden Artikel stellen die Herren Břečka und Geronimus eine interessante asymptotische Formel (14) fest, welche man auf folgende Weise darstellen kann:

$$M \sim \frac{\pi a_1}{n}\left\{\frac{(\nu+1)!}{2^\nu\left(\frac{\nu}{2}\right)!^2}\right\}^2 \quad \begin{pmatrix}\nu = k-1 & \text{für } k \text{ ungerade}\\ \nu = k-2 & \text{für } k \text{ gerade}\end{pmatrix},$$

indem man $L = 2M$, $\varphi = \frac{\pi}{2}$ einführt.

Diese Formel gibt den maximalen Modul M im Intervall $(-1, +1)$ des monotonen Polynoms, dessen Grad n gegen Unendlich strebt, wenn der Koeffizient a_1 von x und auch die von n unabhängigen Werte der Koeffizienten von x^i $(i = 2, 3, 4, \ldots, k)$ vorgelegt sind, wo k eine gegebene endliche Zahl ist.

Die Zahl k geht in diese Formel ganz einfach ein und hat keinen Einfluß auf die Größenordnung von M; die gegebenen Werte der übrigen Koeffizienten spielen dabei keine Rolle.

Wir haben noch keine analoge Formel für den allgemeinen Fall, wenn das Polynom auch nicht-monoton sein kann. Jedoch kann man es für wahrscheinlich halten, daß auch in diesem Falle eine Formel von demselben Charakter gilt. Man könnte dafür einen vollständigen Beweis durch Anwendung der Methode erhalten, welche ich gebraucht habe zum Beweis der Formeln

$$
\text{(I)}\quad \begin{cases} M \sim \dfrac{n^{\frac{l}{2}}}{2^{n-1}\left(\frac{l}{2}\right)!} & (l \text{ gerade}) \\[2ex] M \sim \dfrac{\left(|\sigma_1| + \sqrt{1+\sigma_1^2}\right) n^{\frac{l-1}{2}}}{2^{n-1}\left(\frac{l-1}{2}\right)!} & (l \text{ ungerade}) \end{cases}
$$

für die minimale Abweichung von Null in dem Intervall $(-1, +1)$ des beliebigen Polynoms

$$y(x) = x^n + \sigma_1 x^{n-1} + \sigma_2 x^{n-2} + \ldots + \sigma_l x^{n-l} + \ldots + \sigma_n,$$

wo $\sigma_1, \sigma_2, \ldots, \sigma_l$ bestimmte, von n unabhängige Werte haben [1]).

Die Formeln, die den Formeln (I) für das monotone Polynom analog sind und welche die Formeln der Herren Břečka und Geronimus für den Fall ergänzen, wo die $(l+1)$ ersten Koeffizienten vorgelegt sind, lassen sich folgendermaßen darstellen:

$$
\text{(II)}\quad \begin{cases} M \sim \dfrac{\pi n^{\frac{l}{2}+1}}{2^{n}\left(\frac{l}{2}\right)!} & (l \text{ gerade}), \\[2ex] \left.\begin{aligned} & M \sim \dfrac{\pi n^{\frac{l+1}{2}} (1+\sigma_1^2)}{2^{n}\left(\frac{l-1}{2}\right)!}, \quad |\sigma_1| \leqq 1 \\ & M \sim \dfrac{\pi n^{\frac{l+1}{2}} \sigma_1}{2^{n-1}\left(\frac{l-1}{2}\right)!}, \quad |\sigma_1| \geqq 1 \end{aligned}\right\} & (l \text{ ungerade}). \end{cases}
$$

Die Formel (II) war für $l = 0$ schon von Tchebycheff gegeben worden [2]) und wurde für $l = 1$ unlängst von Herrn Geronimus bewiesen [3]). Der Beweis der allgemeinen Formel (II), den ich für ein beliebiges l gefunden habe, ist noch nicht veröffentlicht worden.

Es ist interessant hinzuzufügen, daß die Formeln, die den einfachen Formeln von W. Markoff [4]) analog sind (welche die minimale Abweichung

---

[1]) S. Bernstein, Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle, p. 24.

[2]) Tchebycheff, Sur les fonctions qui diffèrent le moins possible de zéro, Journal des Mathématiques pures et appliquées, (2) **19** (1874), p. 319–346.

[3]) J. Geronimus, Sur le polynome multiplement monotone qui s'écarte le moins de zéro dont les deux premiers coefficients sont donnés, Comptes Rendus de l'Académie des Sciences de l'U.R.S.S. (1928), p. 485.

[4]) W. Markoff, Über Polynome, die in einem gegebenen Intervalle möglichst wenig von Null abweichen, Math. Annalen **77** (1916), S. 213–258.

von Null des beliebigen Polynoms n-ten Grades geben, wenn der Koeffizient von x^k $[k < n]$ gegeben ist), für monotone Polynome bis jetzt noch nicht bekannt sind.

Herr Geronimus hat nur folgende Formel festgestellt:

Wenn der Koeffizient σ_l von x^{n-l} (wenn n ins Unendliche strebt und l endlich ist) gegeben ist:

$$M \sim \frac{\pi\,|\sigma_l|\,\nu!}{2^{n-l}\cdot n^{\nu-1}} \qquad \left(\nu = \left[\frac{l}{2}\right]\right)^{5)}.$$

---

[5]) J. Gueronimus, Sur le polynome multiplement monotone qui s'écarte le moins de zéro dont un coefficient est donné, Bulletin de l'Académie des Sciences de l'U.R.S.S. (1929), Classe des Sciences Physico-Mathématiques, No. 4, p. 389.

(Eingegangen am 30. 3. 1929.)

Zum Hilbertschen Nullstellensatz.

Von

J. L. Rabinowitsch in Moskau.

Satz. *Verschwindet das Polynom $f(x_1, x_2, \dots, x_n)$ in allen Nullstellen — im algebraisch abgeschlossenen Körper — eines Polynomideals* $\mathfrak{a}$, *so gibt es eine Potenz f^ϱ von f, die zu $\mathfrak{a}$ gehört.*

Beweis. Es sei $\mathfrak{a} = (f_1, f_2, \dots, f_r)$, wo f_i die Variablen $x_1, \dots, x_n$ enthalten. x_0 sei eine Hilfsvariable. Wir bilden das Ideal $\bar{\mathfrak{a}} = (f_1, f_2, \dots, f_r, x_0 f - 1)$. Da der Voraussetzung nach $f = 0$ ist, sobald alle f_i verschwinden, so hat das Ideal $\bar{\mathfrak{a}}$ keine Nullstellen.

Folglich muß $\bar{\mathfrak{a}}$ mit dem Einheitsideal zusammenfallen. (Vgl. etwa bei K. Hentzelt, „Eigentliche Eliminationstheorie", § 6, Math. Annalen 88 [1]).) Ist also $1 = \sum_{i=1}^{i=r} F_i(x_0, x_1, \dots, x_n) f_i + F_0 \cdot (x_0 f - 1)$ und setzen wir in dieser Identität $x_0 = \frac{1}{f}$, so ergibt sich:

$$1 = \sum_{i=1}^{i=r} F_i\left(\frac{1}{f}, x_1, \dots, x_n\right) f_i = \frac{\sum_{i=1}^{i=r} \bar{F}_i f_i}{f^\varrho}.$$

Folglich ist $f^\varrho \equiv 0\,(\mathfrak{a})$, w. z. b. w.

---

[1]) Folgt auch schon aus der Kroneckerschen Eliminationstheorie.

(Eingegangen am 8. 5. 1929.)

Bemerkung zum Hilbertschen Irreduzibilitätssatz.

Von

Karl Dörge in Köln.

Der Hilbertsche Irreduzibilitätssatz[1]) besagt das Folgende: $f(x, t)$ sei ein Polynom von x und t mit rationalen Koeffizienten. Ist dann f als Polynom der beiden Veränderlichen x, t im natürlichen Rationalitätsbereich P irreduzibel[2]), so kann man t als ganze rationale Zahl so spezialisieren, daß f, dann als Polynom von x allein, auch in P irreduzibel wird. Genauer besagt der Satz, daß man t auf unbeschränkt viele Weisen in dieser Art spezialisieren kann.

Aus dem in der Abhandlung von André Weil, Acta mathematica Bd. 52, S. 315 mitgeteilten Satze folgt, wie mir der Entdecker desselben freundlicherweise mitgeteilt hat, das außerordentlich weitgehende allgemeine Resultat: Ein Polynom $f(x, t)$ mit rationalen Koeffizienten zerfällt höchstens dann für unendlich viele ganzzahlige Werte t, wenn es möglich ist, für t einen Ausdruck in einer neuen Unbestimmten u

$$t = c_{-m} u^{-m} + c_{-m+1} u^{-m+1} + \ldots + c_0 + c_1 u + \ldots + c_m u^m \text{ [3])}$$

zu finden, so daß u in t wirklich vorkommt und nach der Substitution f als rationale Funktion von x und u in P identisch zerfällt.

Ich behandle hier den Spezialfall, daß in dem nach x geordneten Polynom der höchste und letzte Koeffizient a_0 und a_n nicht von t abhängen, t also nur in den übrigen Koeffizienten $a_1, \ldots, a_{n-1}$ steckt. Ich beweise dann — unter der selbstverständlichen Voraussetzung des H. I.-S. —, indem ich den H. I.-S. wesentlich benutze: f kann höchstens dann für unendlich viele — sogar rationale — t in P zerfallen, wenn ein durch n

---

[1]) Im folgenden abgekürzt als H. I.-S. Der Satz ist von Hilbert in Crelles Journal **110** bewiesen.

[2]) Diese Voraussetzung nenne ich in Zukunft „die Voraussetzung des H. I.-S.".

[3]) Die c brauchen nicht rational zu sein.

allein — es wird $a_0 \neq 0$ vorausgesetzt, so daß n der genaue Grad ist — einfach zu charakterisierendes Polynom $\Delta(a_0, a_1, \ldots, a_n)$ identisch in t verschwindet. Es zeigt sich an Beispielen, daß diese Bedingung nicht ganz fortgelassen werden kann.

Dann behandle ich im zweiten Teil den noch weiter spezialisierten Fall, in dem $f(x, t)$ so entsteht, daß man in dem Ausdruck $a_0 x^n + a_1 x^{n-1} + \ldots + a_n$ einen Koeffizienten, etwa a_ν, durch t ersetzt, während alle übrigen Koeffizienten a feste rationale Zahlen bedeuten. Dann ergibt sich z. B.: Sind ν und n teilerfremd, $a_0 \neq 0$, $a_n \neq 0$, dann wird $f(x, t)$ höchstens für endlich viele rationale Werte t in P reduzibel, während es, wenn n und ν nicht teilerfremd sind, immer ein Gegenbeispiel gibt. Es kann daraus auch gefolgert werden, daß man durch Ungleichungen zwischen den Koeffizienten $a_0, a_1, \ldots, a_n$ in mannigfacher Weise die Irreduzibilität des Polynoms $\sum_{\nu=0}^{n} a_\nu x^{n-\nu}$ erzwingen kann. Sind n und ν nicht teilerfremd, so wird ferner gezeigt, daß die Anzahl der ganzzahligen t unterhalb S, für welche f in P zerfällt, kleiner als konst. $\sqrt{S}$ ist, wo wieder der Exponent $\frac{1}{2}$ von S — jedenfalls für gerade n, ν — nicht verkleinert werden kann.

I.

1.

$$f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu} \qquad (a_0 \neq 0)$$

sei ein Polynom mit Koeffizienten aus irgendeinem Körper K. Dieses heiße in K halbreduzibel, wenn es sich schreiben läßt in der Form

$$(b_0 x^k + b_1 x^{k-1} + \ldots + b_k)(c_0 x^l + c_1 x^{l-1} + \ldots + c_l) \qquad (k > 0,\ l > 0)$$

derart, daß die äußeren Koeffizienten der beiden Faktoren, also die vier Zahlen

$$b_0,\ b_k,\ c_0,\ c_l,$$

zu K gehören. Aus dem Zerfallen eines Polynoms f in K folgt dann das Halbzerfallen von f in K, aber nicht immer umgekehrt. Die n Wurzeln von f bezeichne man mit

$$x_1, x_2, \ldots, x_n.$$

Die notwendige und hinreichende Bedingung für das Halbzerfallen von f in K ist dann offenbar diese: Es gibt ϱ Wurzeln x, $0 < \varrho < n$, deren Produkt $x_{\alpha_1} x_{\alpha_2} \ldots x_{\alpha_\varrho}$ in K rational ist.

Man bilde das Produkt

$$\Pi(x_{\alpha_1} x_{\alpha_2} \ldots x_{\alpha_\varrho} - \delta)$$

mit einer neuen Unbestimmten δ, multipliziert über alle $\binom{n}{\varrho}$ Kombinationen

von je ϱ der n Wurzeln x. Man erhält dann ein wohlbestimmtes Polynom $F_\varrho(\delta)$ von δ und $\frac{a_1}{a_0}, \ldots, \frac{a_n}{a_0}$ mit ganzen rationalen Koeffizienten. Dann setze man

$$F(\delta) = \prod_{\varrho=1}^{n-1} F_\varrho(\delta).$$

Auf diese Weise ist jedem Polynom f ein bestimmtes Polynom $F(\delta)$ der neuen Unbestimmten δ zugeordnet, und die Bedingung für das Halbzerfallen von f in K ist offenbar die: $F(\delta) = 0$ soll eine Wurzel δ in K haben.

Statt des Körpers K behandle man nun den natürlichen Rationalitätsbereich P. Die Koeffizienten a des Polynoms $f(x)$ sind dann rationale Zahlen. Durch Multiplikation mit dem Generalnenner der a erreicht man, daß dieselben ganze rationale Zahlen sind. Dann denke man sich die äußeren Koeffizienten a_0 und a_n fest vorgegeben, die übrigen Koeffizienten $a_1, a_2, \ldots, a_{n-1}$ veränderlich. Die Wurzeln δ der Gleichung $F(\delta) = 0$ sind dann die Produkte $x_{\alpha_1} x_{\alpha_2} \ldots x_{\alpha_\varrho}$ $(0 < \varrho < n)$. Wird ein solches Produkt rational, so ist es jedenfalls nur endlich vieler rationaler Werte fähig, weil bei reduzierten Darstellungen sein Zähler durch die Teiler von a_n, sein Nenner durch die Teiler von a_0 auf endlich viele Werte beschränkt ist. Man fixiere nun irgendeine Vorschrift, die zu jedem Paar von Zahlen a_0, a_n endlich viele rationale Zahlen

$$\delta_1, \delta_2, \ldots, \delta_S$$

bestimmt von der Art, daß unter ihnen alle etwaigen rationalen Wurzeln δ von $F(\delta) = 0$ enthalten sind. Bei vorgegebenen a_0 und a_n ist

$$\prod_{\sigma=1}^{S} F(\delta_\sigma) = R(a_1, a_2, \ldots, a_{n-1})$$

dann ein Polynom von $a_1, a_2, \ldots, a_{n-1}$, welches durch n, a_0 und a_n eindeutig bestimmt ist, und die Bedingung für das Halbzerfallen von $f(x)$ bei fest vorgegebenem a_0 und a_n ist

$$R(a_1, a_2, \ldots, a_{n-1}) = 0,$$

also eine algebraische Bedingung für $a_1, a_2, \ldots, a_{n-1}$. [4])

---

[4]) Etwas allgemeiner ergibt sich folgendes: Man schreibe die Primzahlzerlegungen von a_0 und a_n vor:

$$a_0 = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r}, \qquad a_n = q_1^{\beta_1} q_2^{\beta_2} \ldots q_s^{\beta_s}.$$

Die Bedingung für das Halbzerfallen von $f(x)$ ist dann das Verschwinden eines durch $n, \alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s$ eindeutig bestimmten Polynoms

$$R(a_1, a_2, \ldots, a_{n-1}, p_1, \ldots, p_r, q_1, \ldots, q_s)$$

von $a_1, \ldots, a_{n-1}, p_1, \ldots, p_r, q_1, \ldots, q_s$.

2. Diese Bedingung $R = 0$ des Halbzerfallens ist nach dem Früheren für das Zerfallen des Polynoms $f(x)$ notwendig. Sie ist aber, wenn man von im folgenden zu charakterisierenden Ausnahmefällen absieht, auch für das Zerfallen hinreichend. Um dies zu zeigen, betrachte man neben dem Polynom $f(x)$ aus dem beliebigen Körper K für $0 < \varrho < n$ das Polynom

$$f_\varrho(x) = a_0^{N_\varrho} \Pi (x - x_{\alpha_1} x_{\alpha_2} \dots x_{\alpha_\varrho}),$$

dessen Wurzeln alle $\binom{n}{\varrho}$ Produkte von je ϱ Wurzeln x sind. N_ϱ bestimme man etwa als kleinstmögliche ganze Zahl, für welche die Koeffizienten von f_ϱ ganze rationale Funktionen von $a_0, a_1, \dots, a_n$ werden. Man erhält dann die Polynome $f_1(x), f_2(x), \dots, f_{n-1}(x)$. Es ist offenbar $f_1(x) = f(x)$. Die Diskriminante von $f_\varrho(x)$ sei Δ_ϱ. Ferner sei[5])

$$\Delta = \Delta_1 \Delta_2 \dots \Delta_{n-1}.$$

Δ ist dann ein durch n eindeutig bestimmtes Polynom

$$\Delta(a_0, a_1, \dots, a_n).$$

$f(x)$ soll nun ein spezielles Polynom heißen, wenn Δ verschwindet. *Ist ein Polynom nicht ein spezielles, so folgt, wenn es halb zerfällt, daß es zerfällt.* Zerfällt es nämlich halb, so ist ein Produkt $x_{\alpha_1} x_{\alpha_2} \dots x_{\alpha_\varrho}$ in K rational, es habe den Wert m. Wegen $\Delta \neq 0$ ist dann die Gleichung

$$x_{\alpha_1} x_{\alpha_2} \dots x_{\alpha_\varrho} - m = 0$$

intransitiv, folglich f in K reduzibel.

3. Wir beweisen nach dieser Vorbereitung den

Satz 1. $f(x, t)$ sei ein Polynom der beiden Veränderlichen x, t mit ganzen rationalen Koeffizienten, welches den Voraussetzungen des H. I.-S. genügt. Geordnet nach Potenzen von x hat dann $f(x, t)$ die Form $\sum_{\nu=0}^{n} a_\nu x^{n-\nu}$. Man setze nun speziell voraus, daß a_0 und a_n nicht von t abhängen, daß also dieses nur in den Koeffizienten $a_1, \dots, a_{n-1}$ steckt. Schließlich möge das zu f, aufgefaßt als Polynom allein von x, gehörende Diskriminantenprodukt Δ nicht identisch in t verschwinden. *Dann gibt es nur endlich viele rationale Zahlen τ, die, statt t in $f(x, t)$ eingesetzt, dieses als Polynom von x in P reduzibel machen.*

Beim Beweise setzen wir den H. I.-S. voraus. Dann schließen wir so: Nach 1. gehört zu $f(x, t)$, aufgefaßt als Polynom von x, ein bestimmtes Polynom

$$R(a_1, a_2, \dots, a_{n-1}) = R^*(t)$$

und nach 2. ein bestimmtes Diskriminantenprodukt

$$\Delta(a_0, a_1, \dots, a_n) = \Delta^*(t).$$

[5]) Es genügt auch, hier $\Delta = \Delta_2 \Delta_3 \dots \Delta_{n-1}$ zu setzen.

Nach Voraussetzung verschwindet $\Delta^*(t)$ nicht identisch. Es hat daher nur endlich viele Nullstellen. Die etwaigen rationalen Nullstellen nenne man $\bar{\tau}$. Die $\bar{\tau}$ sind dann unter den rationalen Zahlen τ alle, welche, in f statt t eingesetzt, dieses zu einem speziellen Polynom — der Veränderlichen x — machen. Für ein rationales τ, das nicht ein $\bar{\tau}$ ist, folgt also aus dem Halbzerfallen von f das Zerfallen in P. Daher kann $R^*(t)$ nicht identisch verschwinden. Denn dann wäre f für jedes rationale τ außer den endlich vielen $\bar{\tau}$ in P reduzibel, was dem H. I.-S. widerspricht. Daher hat $R^*(t)$ nur endlich viele Nullstellen. D. h. aber: f zerfällt nur für endlich viele rationale Zahlen τ halb, also zerfällt es erst recht nur für endlich viele rationale Zahlen τ in P, w. z. b. w.

Handelt es sich um mehrere Parameter t, so ergibt sich ein entsprechender Satz. Durch bekannte, von Kronecker herrührende Verfahren kann man ferner entsprechende Sätze für mehrere Veränderliche x und für beliebige algebraische Zahlkörper zurückführen auf den Fall einer Veränderlichen und den Körper P.

II.

4. Es gilt offenbar der folgende

Hilfssatz. $f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu}$ sei ein Polynom aus dem nicht nur endlich viele Elemente enthaltenden Körper K, $a_0 \neq 0$. *Ist $f(x+h)$ als Polynom von x für hinreichend viele*[6]) *in K rationale Zahlen h halbreduzibel, so ist $f(x)$ in K reduzibel.*

Es muß dann nämlich wenigstens eine Kombination von Wurzeln $x_{\alpha_1}, x_{\alpha_2}, \ldots, x_{\alpha_\varrho}$ $(0 < \varrho < n)$ von $f(x)$ geben derart, daß der Ausdruck

$$(x_{\alpha_1} + h)(x_{\alpha_2} + h) \ldots (x_{\alpha_\varrho} + h)$$

für mehr als ϱ Zahlen h aus K in K rational wird. Dann aber gehören die elementarsymmetrischen Funktionen von $x_{\alpha_1}, \ldots, x_{\alpha_\varrho}$ zu K und dann zerfällt f in K.

Ferner ist es im folgenden bequem, statt mit Polynomen $\sum_{\nu=0}^{n} a_\nu x^{n-\nu}$ mit beliebigen ganzzahligen Koeffizienten aus K zu operieren, sie auf eine Form zu bringen, in welcher der höchste Koeffizient a_0 den Wert 1 hat. Dazu multipliziere man $f(x)$ in bekannter Weise mit a_0^{n-1}, setze dann $a_0 x = x'$ und schreibe schließlich statt der neuen Veränderlichen x' wieder x. Auf diese Weise entspricht dem ganzzahligen Polynom $f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu}$

[6]) Die Anzahl kann allein durch n charakterisiert werden. Es genügt z. B., sie gleich $2^n \cdot n$ zu nehmen.

eindeutig das normierte ganzzahlige Polynom $f^*(x) = \sum_{\nu=1}^{n} a_\nu^* x^{n-\nu}$, wo jetzt aber $a_0^* = 1$, $a_\nu^* = a_0^{\nu-1} a_\nu$ ist. Offenbar zerfällt $f(x)$ dann und nur dann in K halb, wenn $f^*(x)$ in K halb zerfällt.

Nun betrachte man die Polynome, die aus dem Ausdruck $a_0 x^n + a_1 x^{n-1} + \ldots + a_n$ dadurch entstehen, daß man statt eines Koeffizienten[7]) a_ν $(\nu \neq 0)$ die Veränderliche t schreibt, während die übrigen Koeffizienten feste ganze rationale Zahlen sind, jedoch $a_0 \neq 0$, $a_n \neq 0$. Das so entstehende Polynom von x

$$f_\nu(x, t) = a_0 x^n + \ldots + a_{\nu-1} x^{n-\nu+1} + t x^{n-\nu} + a_{\nu+1} x^{n-\nu-1} + \ldots + a_n$$

hat dann in $P(t)$ ganzzahlige Koeffizienten. Ferner ist, wie man leicht erkennt, $f_\nu(x, t)$ in $P(t)$ irreduzibel. Nach dem Hilfssatz gibt es nun eine ganze rationale Zahl h_0 derart, daß

$$f_\nu(x + h_0, t) = \bar{f}_\nu(x, t)$$

als Polynom von x in $P(t)$ nicht halbreduzibel ist. Zu diesem Polynom $\bar{f}_\nu(x, t)$ gehört ferner eindeutig ein normiertes Polynom $\bar{f}_\nu^*(x, t)$ mit ganzen Koeffizienten in $P(t)$, so daß dieses in $P(t)$ nicht halbreduzibel ist. Es gilt dann: $f_\nu(x, t)$ zerfällt für diejenigen rationalen Zahlen t in P, für welche $\bar{f}_\nu^*(x, t)$ in P zerfällt, und nur für diese. Man schreibe $\bar{f}_\nu^*(x, t)$ in der Form $\sum_{\nu=1}^{n} a_\nu^* x^{n-\nu}$, was möglich ist, da es ja noch den Grad n hat. Offenbar ist dann

$$a_n^* = a_0^{n-1} f(h_0).$$

Nun sei S eine ganze positive Zahl und man spezialisiere t in $f_\nu(x, t)$ und $\bar{f}_\nu^*(x, t)$ der Reihe nach auf die Zahlen $1, 2, \ldots, S$. Man erhält dann S Polynome $f_\nu(x, \tau)$ $(\tau = 1, 2, \ldots, S)$ und S Polynome $\bar{f}_\nu^*(x, \tau)$ $(\tau = 1, 2, \ldots, S)$ in P. Unser Ziel ist es, zu untersuchen, wie viele der so erhaltenen Polynome $f_\nu(x, \tau)$ oder, was dasselbe ist, der $\bar{f}_\nu^*(x, \tau)$ in P zerfallen. Für die Zahlen a_n, die jetzt als absolute Glieder in $\bar{f}_\nu^*(x, \tau)$ auftreten, erhält man, da $f(h_0)$ ein linearer Ausdruck von t ist, offenbar die Abschätzung[8])

$$a_n^* < C_1 S.$$

Setzt man daher $C_2 = \sqrt{C_1}$, so ergibt sich: Wenn eines der S Polynome $\bar{f}_\nu^*(x, \tau)$ in P in zwei — wie man ohne Einschränkung annehmen darf,

---

[7]) Den Fall $\nu = 0$ erreicht man, indem man $\nu = n$ setzt und dann das reziproke Polynom nimmt.

[8]) C_1 und die folgenden Konstanten $C_2, \ldots$ lassen sich allein durch n, ν und $a_0, a_1, \ldots, a_{\nu-1}, a_{\nu+1}, \ldots, a_n$ bestimmen. Sie sind also insbesondere von S unabhängig.

normierte — Faktoren zerfällt, so muß das absolute Glied wenigstens eines der beiden Faktoren, welches wegen der Normierung von $\bar{f}_\nu^*$ von selbst eine ganze Zahl ist, $< C_2 \sqrt{S}$ sein.

Nach 1. gehört nun zu $\bar{f}_\nu^*(x, t)$ ein bestimmtes Polynom $F(\delta)$. Für δ setzt man hier der Reihe nach $1, 2, \ldots, [C_2\sqrt{S}], -1, -2, \ldots, -[C_2\sqrt{S}]$ und bilde dann das Produkt $\Pi F(\delta)$. Dies ist ein Polynom

$$\Phi(a_0, a_1, a_2, \ldots, a_{\nu-1}, t, a_{\nu+1}, \ldots, a_n)$$

von t. Der Grad in t kann mit Hilfe der Anzahl der Faktoren des Produktes Π nach oben abgeschätzt werden und ist daher kleiner als $C_3\sqrt{S}$. Φ kann ferner nicht identisch verschwinden, weil daraus folgen würde, daß $\bar{f}_\nu^*(x, t)$ in $P(t)$ halb zerfällt, was ja nicht der Fall war. Daher hat Φ höchstens $C_3\sqrt{S}$ Nullstellen t, und das heißt: Unter den S Polynomen

$$f_\nu(x, \tau) \qquad (\tau = 1, 2, \ldots, S)$$

sind höchstens $C_3\sqrt{S}$ reduzibel.

In ganz entsprechender Weise erhält man dieselbe Abschätzung für die reduziblen unter den Polynomen $f_\nu(x, \tau)$ $[\tau = -1, -2, \ldots, -S]$. Aus diesen beiden Resultaten ergibt sich schließlich:

Satz 2. *Die Anzahl der reduziblen unter den* $2S + 1$ *Polynomen*

$$f_\nu(x, \tau) \qquad (-S \leqq \tau \leqq S)$$

ist $< C_4\sqrt{S}$.[9])

Zusatz: Offenbar gilt, wie jetzt nachträglich folgt, Satz 2 auch, wenn man die Koeffizienten $a_0, \ldots, a_{\nu-1}, a_{\nu+1}, \ldots, a_n$ nur als rational voraussetzt.

Durch Multiplikation mit ihrem Generalnenner, etwa c, geht dann $f_\nu(x, \tau)$ über in $a_0' x^n + \ldots + c\tau x^{n-\nu} + \ldots + a_n'$. Setzt man $c\tau = \tau'$, so wird, wenn τ von $-S$ bis $+S$ variiert, τ' zwischen $-cS$ und $+cS$ liegen. Die Anzahl derjenigen dieser τ', also auch der τ, für welche $f_\nu(x, \tau)$ in P zerfällt, ist dann $< C_4\sqrt{c}\sqrt{S} = C_5\sqrt{S}$.

5. Wie nach Satz 1 zu erwarten ist, läßt sich dies Ergebnis noch weiter verschärfen. Man kann nämlich folgenden Satz beweisen:

Satz 3. Man bilde $f_\nu(x, t)$ für $0 < \nu < n$, $a_0 \neq 0, a_n \neq 0$. *Dieses zerfällt höchstens dann für unendlich viele rationale Werte* τ *in Faktoren der Grade* ϱ *und* σ, *wenn* ϱ *und* σ *Vielfache von* $\frac{n}{d}$ *sind, unter* d *den G. G. T. von* n *und* ν *verstanden*[10]). Sind also z. B. n und ν teilerfremd, so zerfällt $f_\nu(x, t)$ höchstens für endlich viele rationale Zahlen τ.

---

[9]) Dies ist eine wesentliche Verschärfung der Abschätzung, die ich in dieser Zeitschrift **95**, S. 254 angegeben habe.

[10]) Der Satz, für $\nu = 0$ und $\nu = n$ ausgesprochen, ist offenbar selbstverständlich. Er besagt dann nichts.

Soll $f(x, t)$ für unendlich viele rationale τ in Faktoren der Grade ϱ, σ zerfallen, so muß die Gleichung $F_\varrho(\delta) = 0$ bei veränderlichem t durch eine rationale Zahl δ gelöst werden, es muß also $f(x, t)$ im Körper $P(t)$ in Faktoren der Grade ϱ, σ halb zerfallen. Um den Satz zu beweisen, kommt es also allein darauf an, zu zeigen, daß, wenn $f(x, t)$ im Körper $P(t)$ in zwei Faktoren der Grade ϱ, σ halbzerfällt, ϱ und σ Vielfache von $\frac{n}{d}$ sind. Statt von f_ν genügt es, dies von dem normierten Polynom f_ν^* zu zeigen. Die algebraische Funktion x von t, welche durch die Gleichung $f_\nu^*(x, t) = 0$ definiert wird, zerfällt in n Zweige. Jeder der n Zweige wird für hinreichend große t durch eine nach oben abbrechende Laurentreihe einer Wurzel $t^{\frac{1}{q}}$ dargestellt. In unserem Falle müssen, wie man leicht sieht, $n - \nu$ Zweige als höchste Potenz $t^{\frac{1}{n-\nu}}$ und ν Zweige $t^{-\frac{1}{\nu}}$ enthalten. Soll nun f_ν^* in $P(t)$ in zwei Faktoren der Grade ϱ und σ halb zerfallen, so dürfen diese Faktoren als normiert angenommen werden. Dann müssen ihre absoluten Glieder ganze Elemente aus $P(t)$ sein. Die Wurzeln des Faktors ϱ-ten Grades verteilen sich auf ϱ Zweige der algebraischen Funktion, es seien dies $\varkappa$ Zweige, deren Laurentreihe mit $t^{\frac{1}{n-\nu}}$, und λ Zweige, deren Laurentreihe mit $t^{-\frac{1}{\nu}}$ beginnt. Das absolute Glied des Faktors ϱ-ten Grades ist dann eine Laurentreihe einer Wurzel von t, welche den höchsten Exponenten

$$\frac{\varkappa}{n-\nu} - \frac{\lambda}{\nu}$$

hat. Das absolute Glied des andern Faktors beginnt mit dem höchsten Exponenten

$$-\left(\frac{\varkappa}{n-\nu} - \frac{\lambda}{\nu}\right),$$

da das Produkt der beiden absoluten Glieder a_n ist. Da die beiden absoluten Glieder Elemente aus $P(t)$ sein sollten, andererseits wegen der Normierung der Faktoren von selbst ganze Elemente, also Polynome von t sind, folgt $\frac{\varkappa}{n-\nu} - \frac{\lambda}{\nu} = 0$, also $\frac{\varkappa}{\lambda} = \frac{n-\nu}{\nu}$. Ist d der G. G. T von n und ν, also auch der G. G. T. von $n - \nu$ und ν, so folgt jetzt: $\varkappa$ ist Vielfaches von $\frac{n-\nu}{d}$, etwa $\varkappa = \frac{n-\nu}{d} l$. Dann ist $\lambda = \frac{\nu}{d} l$, also $\varkappa + \lambda = \varrho$ und damit auch σ Vielfaches von $\frac{n}{d}$, w. z. b. w.

Zusatz zu Satz 3. Ganz wie im Zusatz zu Satz 2 folgt jetzt wieder nachträglich, daß Satz 3 auch dann gilt, wenn man nur verlangt, daß die Koeffizienten $a_0, \ldots, a_{\nu-1}, a_{\nu+1}, \ldots, a_n$ rationale Zahlen sind.

Einem Teil dieses Satzes wollen wir noch eine etwas andere Formulierung geben. Die Koeffizienten von $f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu}$ seien wieder ganze rationale Zahlen, $a_0 \neq 0$, $a_n \neq 0$. Statt mit f operieren wir nun mit f^*. Auf Grund der Betrachtung von 1. gehört dann zu f^* ein bestimmtes Polynom $F(\delta)$. Als Wurzeln von $F(\delta) = 0$ kommen nun nur ganze rationale Zahlen in Betracht, die ihrer absoluten Größe nach durch a_0 und a_n nach oben beschränkt sind. Die Koeffizienten der Glieder von $R(a_1, a_2, \ldots, a_{n-1})$ sind daher jetzt ganze rationale Zahlen, deren Betrag durch a_0 und a_n nach oben beschränkt ist. Nun sei ν zu n teilerfremd. Wenn dann für irgendein System fester ganzer Zahlen $a_0, a_1, \ldots, a_{\nu-1}, a_{\nu+1}, \ldots, a_n$ das Polynom R die Veränderliche a_ν nicht mehr enthält, so kann für dieses Wertsystem R, wie aus dem Beweise von Satz 3 folgt, nicht verschwinden, also $f(x)$ für keinen einzigen Wert a_ν reduzibel werden. Für jedes andere System $a_0, \ldots, a_{\nu-1}, a_{\nu+1}, \ldots, a_n$ aber muß in R die Veränderliche a_ν wirklich auftreten. Der Koeffizient der höchsten auftretenden Potenz von a_ν ist dem Betrage nach mindestens 1. Daraus folgt

Satz 4. *$f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu}$ habe ganze rationale Koeffizienten. $a_0 \neq 0, a_n \neq 0$. ν und n seien teilerfremd. Dann lassen sich allein durch n und ν zwei ganze positive Zahlen K und M so bestimmen, daß folgendes gilt: S sei das Maximum der absoluten Beträge $|a_0|, \ldots, |a_{\nu-1}|, |a_{\nu+1}|, \ldots, |a_n|$. Ist dann*

$$|a_\nu| > K \cdot S^M,$$

so ist $-R \neq 0$, also $-f(x)$ in P irreduzibel. Es läßt sich also durch Ungleichungen, die für die Koeffizienten vorgeschrieben werden, die Irreduzibilität von $f(x)$ erzwingen. Außer nach der angegebenen lassen sich auf noch mannigfache andere Weise aus der Tatsache, daß das Polynom R verschwinden muß, Ungleichungen ableiten, aus denen die Irreduzibilität von $f(x)$ folgt.

6. Wir wollen schließlich an Beispielen uns über die Schärfe der bewiesenen Sätze einige Klarheit verschaffen. Sind n und ν irgendwelche vorgegebene, nicht teilerfremde Zahlen, so kann man in P Polynome $f_\nu(x, t)$ angeben, die für unendlich viele — sogar ganze — rationale Werte t zerfallen. Sei nämlich d der G. G. T. von n und ν, $n = dn'$, $\nu = d\nu'$. Man wähle dann irgendein Polynom $\varphi(x)$ vom Grade n', dessen Koeffizienten feste ganze rationale Zahlen sind bis auf den ν-ten Koeffizienten. An dessen Stelle setze man $\sqrt[d]{s}$. Das absolute Glied von $\varphi_{d-1}(x)$ möge nicht verschwinden. $\varphi'(x), \varphi''(x), \ldots, \varphi^{(d-1)}(x)$ mögen aus $\varphi(x)$ hervorgehen, indem man statt $\sqrt[d]{s}$ die $d-1$ algebraisch konjugierten Ausdrücke $\varepsilon \sqrt[d]{s}$ einsetzt,

unter ε die von 1 verschiedenen d-ten Einheitswurzeln verstanden.

$$\varphi(x)\cdot\varphi'(x)\dots\varphi^{(d-1)}(x)=\Phi(x)=\sum_{\nu=0}^{n}a_\nu x^{n-\nu}$$

ist dann ein Polynom n-ten Grades im Körper $P(s)$. Jetzt sind alle Koeffizienten in $\Phi(x)$ bis auf den ν-ten, a_ν, feste ganze rationale Zahlen. a_ν läßt sich in der Form schreiben $s+r=t$, unter r eine ganze rationale Zahl verstanden. $\Phi(x)$ ist dann ein Polynom, wie wir es oben als $f_\nu(x,t)$ bezeichnet haben, das bei festen $a_0,\dots,a_{\nu-1},a_{\nu+1},\dots,a_n$, $a_0\neq 0$, $a_n\neq 0$ für unendlich viele rationale Zahlen t in P reduzibel wird, nämlich für alle Zahlen t, für welche $t-r=s$ eine d-te Potenz einer rationalen Zahl ist. Das Beispiel zeigt, daß in dem Zusatz zu Satz 3 die Bedingung, daß n und ν teilerfremd sind, nicht fortgelassen werden kann. Daraus folgt weiter, daß in Satz 1 die Bedingung: $\Delta(a_0,\dots,a_n)$ soll nicht identisch in t verschwinden, nicht ganz fortgelassen werden kann.

Schließlich sei noch ein Beispiel bez. des Satzes 2 angegeben. Es seien n und ν als gerade Zahlen vorgegeben, $n>\nu$, $n=2n'$, $\nu=2\nu'$. $\varphi(x)$ sei ein Polynom n'-ten Grades mit festen ganzen rationalen Koeffizienten $a_0',\dots,a_{\nu-1}',a_{\nu+1}',\dots,a_n'$. An Stelle von a_ν' setze man $\sqrt{s}$. $\varphi'(x)$ entstehe aus $\varphi(x)$, indem man für $\sqrt{s}$ setzt $-\sqrt{s}$. $\Phi(x)=\varphi(x)\,\varphi'(x)$ ist dann ein Polynom $\sum_{\nu=0}^{n}a_\nu x^{n-\nu}$, und es sind $a_0,\dots,a_{\nu-1},a_{\nu+1},\dots,a_n$ von s unabhängig $a_0\neq 0$, $a_n\neq 0$, $a_\nu=s+r=t$, wo r eine ganze rationale Zahl ist. Also hat $\Phi(x)$ die Form eines $f_\nu(x,t)$. Die Anzahl der reduziblen unter den $f_\nu(x,\tau)$ $(-S<\tau<S)$ ist dann im wesentlichen — d. h. bis auf höchstens endlich viele, deren Anzahl nicht von S abhängt — $2\sqrt{S}$, da wir dann und nur dann ein reduzibles Polynom erhalten, wenn $\tau-r=s$ eine Quadratzahl ist. Das zeigt, daß bei geraden n und ν die Abschätzung des Satzes 2 im Exponenten von S nicht verschärft werden kann.

(Eingegangen am 18. 3. 1929.)

Begründung der projektiven Geometrie im offenen Kontinuum.

Von

Hans Mohrmann in Darmstadt.

Inhaltsübersicht.

Einleitung.

In den letzten fünfzig Jahren hat die Loslösung der Analysis vom Mutterschoß der Geometrie ihr Ende gefunden. Die Trennung des Begriffs der *Zahl* von demjenigen der *meßbaren Größe* (Länge einer Strecke), die noch P. Du Bois-Reymond[1]) für absurd hielt, hat dabei beträchtliche Wehen verursacht. Die Schwierigkeiten, die zu überwinden waren, spiegeln sich vielleicht am sichtbarsten in den Umständen, die die Einführung der Zahl in die Geometrie unabhängig vom Maßbegriff, also, modern gesprochen, ohne Benutzung von Kongruenzaxiomen bereitet hat; oder anders ausgedrückt: die die Begründung der projektiven Geometrie auf Grund der ihr eigentümlichen projektiven Axiome allein verursacht hat. Es gibt wohl kaum einen mathematischen Satz, über den so viel geschrieben worden ist, wie über den Fundamentalsatz der projektiven Geometrie[2]).

[1]) Allgemeine Funktionentheorie, Tübingen 1882. — Vgl. auch Reye, Geometrie der Lage I, 1, 4. Aufl. (1898), S. 8: „Auf die höhere Analysis, dieses mächtige Werkzeug der modernen Mathematik, müssen wir schon deshalb verzichten, weil wir das Maß nicht benutzen."

[2]) Vgl. Hölder, Math. Annalen 65 (1908), S. 161 ff. Die (100 Seiten umfassende) Arbeit Hölders über die „Geometrie der projektiven Geraden" gibt wohl einen vollständigen Überblick über die frühere Literatur. Dabei zeigt sich die Stärke Hölders im Blick für die Schwächen seiner Vorgänger, seine Schwäche im Mangel an Blick für ihre Stärken.

Wir wollen im folgenden zeigen, daß diese Schwierigkeiten nicht nur in der Natur der Sache begründet sind. Dabei folgen wir einem Gedanken von F. Lindemann[3]), dessen große Tragweite allerdings weiterhin nur D. Hilbert[4]) erkannt zu haben scheint, der sich den Lindemannschen Gedanken ausdrücklich zu eigen gemacht und mehrfach verwandt hat[5]).

I.

Axiome des offenen projektiven Kontinuums.

Wir legen das folgende aus sieben „linearen" (einschließlich dem Stetigkeitsaxiom), zwei „ebenen" und einem „räumlichen" Axiom bestehende System zugrunde, das sich ohne Mühe auf Räume von beliebig vielen Dimensionen ausdehnen läßt:

Das *Element* der Punktgeometrie ist der *Punkt*.

I, 1. *Es gibt wenigstens zwei voneinander verschiedene Punkte.*

I, 2. *Zwei voneinander verschiedene Punkte bestimmen eine Menge von Punkten, die wir g-Linie nennen.*

I, 3. *Irgendzwei voneinander verschiedene Punkte einer und derselben g-Linie bestimmen diese g-Linie.*

I, 4. *Wenn A, B, C Punkte einer g-Linie sind und B zwischen A und C liegt, so liegt B auch zwischen C und A.*

I, 5. *Wenn A und C zwei voneinander verschiedene Punkte einer g-Linie sind, so gibt es stets wenigstens einen Punkt B, der zwischen A und C liegt, und wenigstens einen Punkt D, so daß C zwischen A und D liegt.*

I, 6. *Unter irgend drei voneinander verschiedenen Punkten einer g-Linie gibt es stets einen und nur einen, der zwischen den beiden anderen liegt.*

Erklärung. Wir betrachten auf einer g-Linie zwei voneinander verschiedene Punkte A und B; wir nennen das System der beiden Punkte A und B eine *Strecke* und bezeichnen sie mit $\overline{AB}$ oder $\overline{BA}$. Die Punkte zwischen A und B heißen Punkte der Strecke $\overline{AB}$ oder auch *innerhalb* $\overline{AB}$ gelegen; die Punkte A und B heißen *Endpunkte* der Strecke $\overline{AB}$. Alle übrigen Punkte der g-Linie heißen *außerhalb* der Strecke $\overline{AB}$ gelegen.

---

[3]) Clebsch-Lindemann, Vorles. über Geometrie 2 (1891), Teil I, S. 433ff.

[4]) Grundlagen der Geometrie, Anhang I, S. 113 der 6. Aufl.

[5]) Was wohl deshalb so wenig bekannt ist, weil fast immer von den Grundlagen der Geometrie Hilberts gesprochen und nur selten beachtet wird, daß Hilbert verschiedene Eingänge in die Grundlagen gezeigt hat, keineswegs nur einen solchen, bei dem die Stetigkeitsforderungen den Schlußstein des Axiomensystems bilden.

I, 7. (Stetigkeitsaxiom[6]).) *Jedem Punkte einer Strecke entspricht eine reelle Zahl eines Intervalls und umgekehrt.*

Die Zuordnung der Zahlen zu den Punkten der Strecke, d. i. die *Numerierung* der Punkte, erfolgt alsdann auf Grund der Axiome des „*Zwischen*“, wobei noch vielfach-unendlich viele Möglichkeiten bestehen. Die primitivste *stetige Skala* (zum Unterschiede von den *projektiven* und *metrischen*) erhält man, wenn man den Endpunkten A und B der Strecke $\overline{AB}$ die (reellen) Zahlen α und β zuordnet, wobei etwa $\alpha < \beta$ sei, einem beliebigen Punkte C zwischen A und B aber eine *beliebige* Zahl γ, $\alpha < \gamma < \beta$ usw.

I, 8. *Es gibt wenigstens drei nicht auf einer und derselben g-Linie gelegene Punkte.*

I, 9. *Wenn A, B, C irgend drei nicht in einer und derselben g-Linie gelegene Punkte sind, D ein Punkt der Strecke $\overline{BC}$ und E ein Punkt der Strecke $\overline{AD}$ ist, dann gibt es einen der Strecke $\overline{AB}$ angehörigen Punkt F derart, daß E auf der Strecke $\overline{CF}$ liegt*[7]).

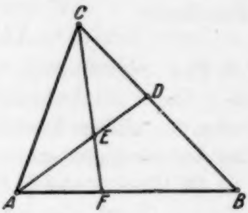


Fig. 1.

Erklärung. Die Menge der Punkte, die durch drei nicht auf einer und derselben g-Linie liegende Punkte A, B, C vermöge den Punkten ihrer Verbindungsstrecken oder g-Linien und denjenigen der Verbindungsstrecken oder g-Linien der Punkte dieser bestimmt ist, heißt *Dreieck* oder *E-Fläche* (*Ebene*).

I, 10. *Es gibt wenigstens vier nicht in einer und derselben E-Fläche gelegene Punkte.*

Tetraeder und *Raum*, *Simplex* und *Raum von mehr als drei Dimensionen* werden ebenso erklärt wie Dreieck und E-Fläche.

Wollen wir uns auf ein dreidimensionales Kontinuum beschränken, so brauchen wir noch ein *Schlußaxiom*:

[6]) Die auf die Vorstellung der *Iteration* gegründete *analysis infinitorum* der projektiven Geometrie (unbegrenzt fortsetzbare Vierseitskonstruktionen) ist keine andere als die auf die Vorstellung der *natürlichen Zahlenreihe* gegründete analysis infinitorum der reellen Zahlen. — Vgl. Weyl, Das Kontinuum, Leipzig 1918, S. 72f. Auch wenn man den Standpunkt Weyls nicht einnimmt (Fréchet, v. Kerékjártó) liegt für uns kein Grund vor, jene analysis noch einmal im geometrischen oder mengentheoretischen Gewande zu begründen.

[7]) Peano, Sui fondamenti di geometria, Rivista di matematica 4 (1894), p. 65. E. H. Moore, Trans. of the Amer. Math. Soc. 3 (1902), p. 147. F. Schur, Grundlagen der Geometrie (1909), S. 7ff.

I, 11. *Außerhalb eines Raumes gibt es keinen Punkt.*

Andernfalls läßt sich unser Axiomensystem ohne weiteres auf Räume von beliebig vielen Dimensionen ausdehnen. In *jedem* Falle können wir noch ein *Vollständigkeitsaxiom* hinzufügen:

I, 12. *Die Elemente der Geometrie bilden ein System von Dingen, das bei Aufrechterhaltung sämtlicher genannten Axiome keiner Erweiterung mehr fähig ist*[8]).

Dieses hat dann allerdings einen völlig anderen Charakter als bei Hilbert, wo es mit dem „Archimedischen Axiom" zusammen die Forderung der Stetigkeit ausmacht, die wir in I, 7 gestellt haben. Es hat bei uns vielmehr den Zweck, unter allen zulässigen („nirgends-konkaven", offenen) Kontinuen das *umfassendste* auszusondern.

Betrachten wir z. B. als Punkte unserer Geometrie, die Punkte der *Kugelhaube*

$$X^2 + Y^2 + Z^2 = 2, \quad Z > 1$$

(X, Y, Z rechtwinklige kartesische Koordinaten im Euklidischen Raum), als g-Linien die Großkreise der Kugel, soweit sie auf der Kugelhaube verlaufen, so erfüllen diese Punkte und g-Linien sämtliche Axiome I, 1—10. Das *Vollständigkeitsaxiom* I, 12 aber ist nicht erfüllt.

Die Punkte und g-Linien der offenen *Halbkugel*

$$X^2 + Y^2 + Z^2 = 2, \quad Z > 0$$

indessen erfüllen — bei Beschränkung auf ein zweidimensionales Kontinuum — auch das Vollständigkeitsaxiom.

Hieraus erkennt man leicht, daß das *Hilbertsche Parallelenaxiom*: *Es sei a eine beliebige g-Linie und A ein Punkt außerhalb a: dann gibt es in der durch a und A bestimmten E-Fläche höchstens eine g-Linie* (und dann eine und nur eine), *die durch A läuft und a nicht schneidet,* auf Grund unserer Axiome I, 1—12 ein *beweisbarer Satz* ist. (Hierin ist wohl auch der Grund zu suchen, warum das Hilbertsche Parallelenaxiom nur in der ersten Auflage der „Grundlagen" vor den Kongruenzaxiomen gestanden hat.) In der Tat ist das Hilbertsche Parallelenaxiom dem Euklidischen nur deshalb äquivalent, weil die Axiome des „Zwischen", die Euklid fremd waren, insbesondere Hilberts II, 3 (unser I, 6) in Verbindung mit Hilberts erstem Kongruenzaxiom (III, 1) vom Streckenabtragen die *unendliche Länge* der g-Linie fordern (die Euklid erst im Parallelenaxiom voraussetzt) und damit die *elliptische Geometrie* (auch im offenen Kontinuum) ausschließen[9]).

[8]) Hilbert, Grundlagen, § 8.

[9]) Siehe hierüber mein demnächst im Verlag der „Akademischen Verlagsgesellschaft" in Leipzig erscheinendes Büchlein: Einführung in die Nicht-Euklidische Geometrie.

II.

Die Beweismittel.

1. Wir dürfen hier als bekannt voraussetzen[9]), daß man auf Grund unserer Axiome I, 1—6, 8—10, also ohne Benutzung des Stetigkeitsaxioms I, 7 den folgenden *Satz vom 4. harmonischen Punkte* beweisen kann:

a) *Sind A, B, C drei voneinander verschiedene Punkte einer g-Linie g derart, daß B zwischen A und C liegt, so gibt es auf g zu C in bezug auf A und B einen und nur einen 4. harmonischen Punkt D.* Dabei heißen nach v. Staudt[10]) vier Punkte A, B, C, D einer und derselben g-Linie harmonisch, wenn A und B Schnittpunkte der Gegenseitenpaare eines eigentlichen Vierecks sind und C und D auf den Diagonalen dieses Vierecks liegen. D liegt dann notwendig zwischen A und B.

Wir dürfen hier ferner den folgenden leicht auf Grund derselben Axiome elementar beweisbaren Satz über *harmonische Strahlen voraussetzen*[11]):

b) *Werden die g-Linien a, b, c, d, die vier hormonische Punkte A, B, C, D einer g-Linie g aus einem Punkte S außerhalb g projizieren, von einer beliebigen weiteren g-Linie g_1 in vier Punkten $A_1\, B_1\, C_1\, D_1$ geschnitten derart, daß S nicht zwischen A und A_1, B und B_1, C und C_1, D und D_1 liegt, so sind auch $A_1 B_1 C_1 D_1$ vier harmonische Punkte. Die g-Linien $a\; b\; c\; d$ heißen* (deshalb) *harmonische Strahlen.*

2. Bezeichnet man mit Pringsheim[12]) als *Systembruch* mit der Basis b (b positiv ganz $\geqq 2$) einen Ausdruck von der Form

$$\sigma_\nu = g + \frac{a_1}{b} + \frac{a_2}{b^2} + \ldots + \frac{a_\nu}{b^\nu},$$

wo g eine ganze Zahl, auch die Null, vorstellen kann und die a „Ziffern" der Zahlenwerte $0, 1, 2, \ldots, b-1$ bedeuten, so gilt der (uns aus der Dezimalzahl-Theorie geläufige) *grundlegende Satz*:

Jede reelle Zahl kann durch einen und nur einen unbegrenzt fortsetzbaren Systembruch[13]) *mit vorgeschriebener Basis b*

---

[10]) Geometrie der Lage, Nürnberg 1847.

[11]) Im allgemeinen zerstört Projektion das „Zwischen". — Wir brauchen die Sätze a) und b) nur in dieser engen Form.

[12]) Vorlesungen über Zahlen- und Funktionenlehre 1, 1 (1916), S. 111 ff.

[13]) Allenfalls mit der eingliedrigen Periode $b-1$;

für $b = 10$ ist z. B. $\frac{1}{2} \equiv 0{,}5 = 0{,}4999\ldots$ und

für $b = \;2$ ist $\frac{1}{2} \equiv 0{,}1 = 0{,}0111\ldots$.

$$\sigma = g + \frac{a_1}{b} + \frac{a_2}{b^2} + \dots + \frac{a_\nu}{b^\nu} + \dots$$

dargestellt werden, und umgekehrt stellt jeder solche Systembruch eine und nur eine reelle Zahl dar.

Es ist also

$$\sigma_\nu < \sigma \leqq \sigma_\nu + \frac{1}{b^\nu} \quad (\text{für jedes } \nu = 0, 1, 2, \dots).$$

Damit haben wir die Beweismittel zusammengestellt.

III.

Über den Fundamentalsatz der projektiven Geometrie.

1. Seien wieder X, Y, Z gewöhnliche rechtwinklige kartesische Koordinaten im Euklidischen Raum. Alsdann stellen die Gleichungen und Ungleichungen

$$\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1 \qquad (0 < a < b < c;\ X \geqq 0,\ Y \geqq 0,\ Z \geqq 0)$$

einen *Oktanten eines dreiachsigen Ellipsoids* dar. Die *inneren* Punkte dieses Oktanten ($X > 0$, $Y > 0$, $Z > 0$) betrachten wir als Punkte einer Geometrie, die geodätischen (kürzesten) Linien des Ellipsoid-Oktanten als ihre g-Linien. Das System dieser Punkte und g-Linien genügt unseren sämtlichen Axiomen I, 1—9. Dem *„räumlichen"* Axiom I, 10 indessen genügt es, nach einem bekannten Satze von Beltrami[14]), nicht, d. h. *ein den Axiomen* I, 1—9 *genügendes System von Punkten und g-Linien kann im allgemeinen nicht als Teil eines entsprechenden räumlichen Systems von Punkten und g-Linien aufgefaßt werden.*

Wir wollen zeigen, daß *eine den Axiomen* I, 1—7 *genügende g-Linie immer als Teil eines räumlichen Systems betrachtet werden kann*, oder mit anderen Worten, daß eine allgemeine *stetige Skala* (Abschnitt I dieser Arbeit) immer als *projektive Skala* aufgefaßt werden kann.

Die drei Punkte, in denen die positiven Koordinatenachsen unser Ellipsoid durchstoßen, bezeichnen wir mit X, Y und O. Wir *numerieren* die Randpunkte unseres Ellipsoid-Oktanten auf OX und OY, d. h. wir stellen zwei beliebige *stetige Skalen* OX und OY her, indem wir dem Punkte O den Wert Null, den Punkten X und Y je den Wert ∞ zuweisen und irgendeinem Punkte zwischen O und X bzw. O und Y den Wert x bzw. y $0 < x < \infty$, $0 < y < \infty$, usw. Alsdann sind sowohl die

[14]) Annali di matematica 7 (1865), S. 187.

Punkte x auf OX zwischen O und X als auch die Punkte y auf OY zwischen O und Y den positiven Zahlen zwischen 0 und ∞ eindeutig umkehrbar zugeordnet. Die g-Linien, welche die Punkte der X-Achse aus Y projizieren, wählen wir als Koordinatenlinien $x =$ konst., diejenigen, welche die Punkte der Y-Achse aus X projizieren, als Koordinatenlinien $y =$ konst.

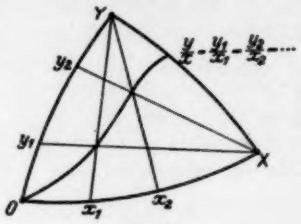


Fig. 2.

Durch jeden Punkt im Innern des Ellipsoid-Oktanten geht eine und nur eine Koordinatenlinie des einen und des anderen Systems.

Wir bezeichnen die Punktmengen im Innern des Ellipsoid-Oktanten, deren Koordinaten eine Gleichung von der Form

$$Ax + By + C = 0$$

erfüllen, als γ-Linien. Alsdann genügen die Punkte des Ellipsoid-Oktanten und seine γ-Linien unseren sämtlichen Axiomen I, 1—10. Und wir könnten, wenn wir uns paradox ausdrücken wollten, sagen, wir hätten die projektive Geometrie mit alleiniger Benutzung linearer und ebener projektiver Axiome begründet.

Indessen sind nur die Koordinatenlinien

$$x = \text{konst.} \quad \text{und} \quad y = \text{konst.}$$

zugleich g- und γ-Linien. Hätten wir nur *eine* Skala *direkt* hergestellt, etwa die auf der X-Achse, und diese dann zunächst aus Y auf eine beliebige g-Linie g durch O projiziert, die so gewonnene Skala wiederum aus X auf die Y-Achse, so wäre auch die γ-Linie

$$y = x$$

mit einer g-Linie identisch, weitere γ-Linien indessen nicht. Hieran würde sich auch dann noch nichts ändern, wenn $a = b = c > 0$, d. h. unser Ellipsoid eine Kugel wäre, obwohl dann die g-Linien des *Kugel-Oktanten* (und seine Punkte) unseren sämtlichen Axiomen I, 1—10 genügen, so daß sowohl für die γ-Linien als auch für die g-Linien des Kugel-Oktanten die Sätze der ebenen projektiven Geometrie (im offenen Kontinuum) gültig wären.

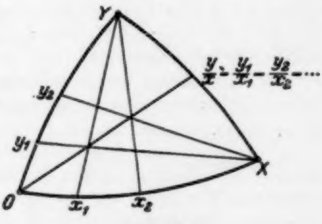


Fig. 3.

Vier Punkte der X-Achse, die in bezug auf die γ-Linien *harmonisch* sind, sind es in bezug auf die g-Linien (im allgemeinen) *nicht*. *Eine Los-*

lösung der projektiven Geometrie der Geraden von der Ebene hat also keinen geometrischen Sinn[15]).

2. Wir ersehen daraus weiter, daß man, um den Fundamentalsatz der projektiven Geometrie auf Grund unserer Axiome I, 1—10 zu erweisen, die Koordinaten nicht willkürlich einführen darf. Dies hat vielmehr so zu geschehen — und das ist der in der Einleitung erwähnte *Lindemannsche Gedanke* —, daß das System der γ-Linien mit dem System der g-Linien zusammenfällt. Kann man das zeigen, so ist damit auch der Fundamentalsatz der projektiven Geometrie erwiesen, sowie die Gültigkeit des *Pascalschen Satzes*[16]) (im offenen Kontinuum) usw.

Notwendige Voraussetzung ist demnach, daß vier Punkte einer Koordinatenachse, die in bezug auf die g-Linien harmonisch sind, auch in bezug auf die γ-Linien harmonisch liegen. Hieraus erklärt es sich, daß man eine im v. Staudtschen geometrischen Sinne *projektive Skala* im Gegensatz zur *stetigen* (im Hölderschen Sinne projektiven) *Skala nur mit Hilfe von Vierseits-Konstruktionen herstellen* kann, indem man drei Punkten A, B, C beliebige Zahlen x_1, x_2, x_3 zuweist, dem vierten harmonischen Punkt D hingegen diejenige Zahl x, die sich aus der Gleichung

$$(ABCD) = (x_1 x_2 x_3 x) = -1$$

ergibt, wo $(x_1 x_2 x_3 x_4)$ das *Doppelverhältnis* der Elemente $x_1 x_2 x_3 x_4$

---

[15]) Hölder (a. a. O., S. 248) nennt eine ein-eindeutige Beziehung zwischen den sämtlichen Punkten einer Geraden g und den sämtlichen Punkten einer Geraden g' (die Geraden können auch koinzidieren) projektiv, wenn die Bedingung erfüllt ist, daß vier harmonischen Punkten von g stets vier harmonische Punkte von g' entsprechen und umgekehrt. Hölder konstatiert ausdrücklich die Übereinstimmung seiner Definition der Projektivität mit derjenigen v. Staudts und fügt hinzu, daß die Definition durch (mehrmaliges) Projizieren für seine Zwecke unbrauchbar sei. Da Hölder trotz den 100 Seiten, die er der „Geometrie der projektiven Geraden" widmet, keine Gelegenheit findet, zu erklären, was man unter vier harmonischen Punkten zu verstehen habe, während er auf der anderen Seite als *ein* Ziel seiner Arbeit die „*Loslösung der projektiven Geometrie der Geraden von der Ebene*" bezeichnet, so kann die Meinung entstehen, es sollten an die Stelle der v. Staudtschen Definition von vier harmonischen Punkten die hierauf bezüglichen *Hölderschen Axiome* III bis VI treten, um so mehr, als Hölder sagt (S. 167): „Dabei ist es aber nun notwendig, daß außer den für die Gerade gültigen Axiomen der Anordnung und ... der Stetigkeit gewisse die harmonischen Punkte betreffende Tatsachen vorausgesetzt werden, die man bei einem anderen Ausgangspunkt der Untersuchung aus den Tatsachen der Ebene bzw. des Raumes beweist."

[16]) Der passend spezialisierte „*Pascalsche Satz*" spielt bekanntlich bei der Begründung der projektiven Geometrie mit Hilfe der Kongruenzaxiome eine wichtige Rolle; er ist dem Fundamentalsatz äquivalent. Vgl. z. B. Schur, Grundlagen, a. a. O. S. 170.

$$\frac{x_1 - x_3}{x_1 - x_4} : \frac{x_2 - x_3}{x_2 - x_4},$$

die bekannte *Invariante* gegenüber linearer Substitution, bedeutet.

Man ersieht daraus aber auch, daß der *Fundamentalsatz der projektiven Geometrie kein linearer Satz* ist, obwohl man ihn scheinbar so aussprechen kann, indem man mit v. Staudt die Projektivität nicht durch Perspektivität, sondern durch das Entsprechen harmonischer Gebilde erklärt und vergißt, daß harmonische Punktquadrupel nur mit Hilfe der Ebene geometrisch definiert werden können.

Wir können den Sachverhalt auch so ausdrücken: Die projektive Geometrie der Ebene ist das Primäre, der sogenannte Fundamentalsatz der projektiven Geometrie das Sekundäre. Wir geben ihm die folgende Fassung:

Fundamentalsatz der projektiven Geometrie (im offenen Kontinuum). *Numeriert man mittels wiederholter (unbegrenzt fortsetzbarer) Vierseits-Konstruktionen die Punkte einer Strecke AC, indem man etwa den Endpunkten A und C der Strecke die Werte* 0 *und* ∞, *einem beliebigen Punkte B zwischen A und C die Zahl* 1 *zuordnet, so ist die entstehende projektive Skala vom Wege der Herstellung unabhängig.*

Damit treten aber auch schon die eigenartigen Schwierigkeiten zutage, die bei einer Begründung der projektiven Geometrie auf Grund der ihr eigentümlichen projektiven Axiome allein zu überwinden sind. Wir erwähnen nur zwei: 1. Es ist zu zeigen, daß vier Punkte, für deren Koordinaten das Doppelverhältnis

$$(x_1 x_2 x_3 x_4) = -1$$

ist, ein (im geometrischen Sinne) harmonisches Quadrupel bilden; 2. Es ist zu zeigen, daß Punkte mit gleichen Koordinaten, die sich auf verschiedenen Wegen ergeben haben, identisch sind.

Dem letzten Umstande können wir sofort entnehmen, daß der Fundamentalsatz der projektiven Geometrie, auch unter Voraussetzung des Stetigkeitsaxioms I, 7, d. h. *unter Ausschluß späterer Hinzunahme sogenannter Nicht-Archimedischer Elemente*[17]), nicht ohne Stetigkeitsbetrachtungen erwiesen werden kann[18]). Man braucht nur zu beachten, daß den Vierseits-Konstruktionen gegenüber ein und dasselbe Element *„rational“ und „irrational“* sein kann. Beispielsweise ergibt sich $\frac{1}{3}$ aus 0, 1, ∞ mit

[17]) Siehe Hilbert, Grundlagen, § 12. — Eine *angebliche* Begründung der (projektiven) Geometrie mit Hilfe von Kongruenzaxiomen *ohne* Stetigkeitsaxiom hat nur unter dieser Voraussetzung einen Sinn, d. h. die so entstehende Geometrie ist *unfertige, nicht* etwa *elementare* Geometrie.

[18]) Vgl. Hilbert, Grundlagen, Kap. VI.

zwei Schritten, nämlich so:

$$(0 \quad 1 \quad \infty \quad x_1) = -1 \qquad \left(x_1 = \frac{1}{2}\right);$$

$$\left(0 \quad \frac{1}{2} \quad 1 \quad x_2\right) = -1 \qquad \left(x_2 = \frac{1}{3}\right).$$

Wendet man dagegen den Prozeß des fortgesetzten „Halbierens" an:

$$(x_1 \quad x_2 \quad \infty \quad x) = -1 \qquad \left(x = \frac{x_1 + x_2}{2}\right),$$

so kann man $\frac{1}{3}$ nicht mit endlich vielen Schritten erreichen, da es *nur eine* Entwicklung von $\frac{1}{3}$ in einen (unbegrenzt fortsetzbaren, periodischen) *dyadischen Bruch* gibt (II, 2):

$$\frac{1}{3} = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \ldots$$

oder, in dyadischer Schreibweise, wegen

$$1 = 2^0 \equiv 1{,}000\ldots \quad \text{und} \quad 3 = 2^1 + 2^0 \equiv 11{,}000\ldots,$$

$$\frac{1}{3} \equiv \frac{1}{11} = 0{,}010101\ldots.$$

IV.

Beweis des Fundamentalsatzes.

Wir schreiten nunmehr zum Beweise des Fundamentalsatzes der projektiven Geometrie im *begrenzten Gebiete* oder, was dasselbe besagt, im offenen Kontinuum auf Grund unserer Axiome I, 1—10. Als Kontinuum wählen wir das Innere eines Dreiecks (YOX). Man sieht dann leicht, daß man diesem Kontinuum, ohne mit den vorausgesetzten Axiomen in Widerstreit zu geraten, noch drei weitere, völlig gleichartige „Quadranten", samt den ihnen gemeinsamen Randpunkten, aber mit Ausschluß der Punkte $X^{\pm\infty}$ und $Y^{\pm\infty}$, nämlich

$$(YO-X), \quad (-XO-Y), \quad (-YOX)$$

hinzufügen kann. Und daß man in diesem *umfassendsten* zweidimensionalen, offenen Kontinuum (im Sinne des Vollständigkeits-Axioms I, 12) wiederum Kontinua abgrenzen kann, für deren Inneres ebenfalls sämtliche Axiome I, 1—10 gültig sind.

Hierzu ist zunächst nötig, mit Hilfe von Vierseits-Konstruktionen eine projektive X-Skala OX herzustellen. Das kann auf unendlich viele Weisen geschehen, und es ist nach dem Fundamentalsatze gleichgültig, auf welchem Wege es geschieht. Aber es ist *nicht gleichgültig für den Beweis.* Wollen wir unser Ziel erreichen, so müssen wir die schon erwähnten Schwierigkeiten

berücksichtigen. Hierzu ist in erster Linie erforderlich, die Skala so anzulegen, daß wir sicher sind, jeden Punkt der Trägerstrecke ein und nur einmal zu *numerieren*.

Dies geschieht am einfachsten dadurch[19]), daß man den Endpunkten der Skalenträgerstrecke OX die Werte 0 und ∞ zuweist, und einem beliebigen Punkte zwischen O und X die Zahl 1. Dann sucht man die Punkte mit ganzzahligen Abszissen auf; und endlich unterteilt man jede Strecke, deren Endpunkte zwei aufeinanderfolgende ganze Zahlen zu Abszissen haben, auf die gleiche Weise, durch fortgesetztes „Halbieren".

Zu diesem Zwecke wenden wir den früher erklärten Prozeß der Konstruktion des *vierten harmonischen* Punktes D (zu einem Punkte C in bezug auf zwei von diesem und voneinander verschiedene Punkte A und B, von denen B zwischen A und C liegt) auf die beiden folgenden Weisen an:

(1) $$(ABCD) = (\infty\, x_2\, x_1\, x) = -1,$$

d. h.

$$x - x_2 = x_2 - x_1$$

(Antragen der gleichen Strecke einer metrischen Skala im Euklidischen Bilde), woraus sich, wenn man mit $x_1 = 0$ und $x_2 = 1$ beginnt, die Reihe der *natürlichen* (positiven, ganzen) *Zahlen* ergibt[20]);

(2) $$(ABCD) = (x_1\, x_2\, \infty\, x) = -1,$$

d. h.

$$x = \frac{x_1 + x_2}{2}$$

(Halbieren einer Strecke einer metrischen Skala im Euklidischen Bilde), woraus sich die erwähnten Unterteilungen ergeben.

Wir wollen nun unsre „Abszissen"-Skala x soweit fertiggestellt denken, daß die Differenz irgend zweier aufeinanderfolgender Abszissen $\frac{1}{2^k}$ (k pos. ganz) beträgt. Alsdann projizieren wir die Skala OX aus Y auf eine beliebige g-Linie OU im Innern des Dreiecks (OXY), und die so entstehende Skala aus X auf OY, wodurch wir eine „Ordinaten"-Skala y erhalten. Die g-Linien durch Y und X wählen wir als Koordinatenlinien

$$x = \text{konst.} \quad \text{und} \quad y = \text{konst.}$$

[19]) Vgl. Clebsch-Lindemann a. a. O., S. 436.

[20]) Nur der Mathematiker, an dessen axiomatischen Gebilden noch das (den geometrischen Gebilden der Wirklichkeit eigentümliche) deiktische Begriffsmerkmal des Ausgedehntseins haftet, wittert hier das Fehlen eines weiteren (Archimedischen) Axioms: die „Linien" der „axiomatischen" Geometrie sind aber, ebenso wie diejenigen der „analytischen", nur Mengen von Elementen, die wir Punkte oder Zahlen nennen; sonst nichts.

Und die Schnittpunkte der Koordinatenlinien bezeichnen wir als *Gitterpunkte*, und wir nennen

$$\frac{1}{2^k}$$

die *Maschenweite des* (unbegrenzt fortsetzbaren) *offenen Gitters.*

Weiter bezeichnen wir, wie früher, als γ-Linien solche Punktmengen unsres Kontinuums, deren Koordinaten (x, y) linearen Gleichungen

$$Ax + By + C = 0$$

genügen. Alsdann wollen wir den folgenden, für uns *grundlegenden Satz* beweisen:

Irgend drei Punkte eines offenen Gitters, die auf einer und derselben γ-Linie liegen, gehören auch einer und derselben g-Linie an.

Zu diesem Zwecke brauchen wir nur die Figuren (4) und (5) zu betrachten, von denen (5) das metrische Euklidische Bild von (4) ist.

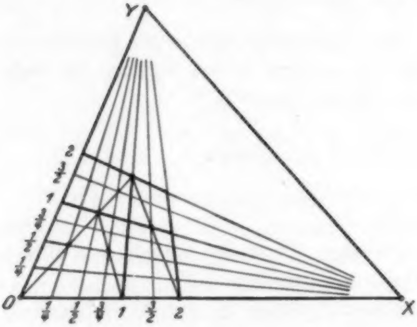


Fig. 4.

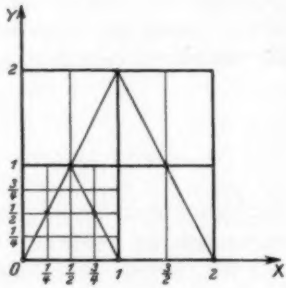


Fig. 5.

Auf Grund der Skalenkonstruktion und des Satzes (II, 1, b), nach dem vier harmonische Punkte durch vier harmonische Strahlen projiziert werden und, was dasselbe besagt, vier harmonische Strahlen nach vier harmonischen Punkten geschnitten werden, erkennt man sofort die Richtigkeit unsres Satzes für die γ-Linien $y = 2x$ und (ihre Spiegelbilder)

$$y = -2(x-1) \quad \text{sowie} \quad y = -2(x-2).$$

Hierbei kommen als projizierende Strahlen nur *Koordinatenlinien* in Betracht. Wählt man aber jetzt die Punkte $(x = \frac{1}{2},\ y = 1)$ und $(x = 1, y = 2)$ zu Projektionszentren, so erkennt man, daß auch die Punkte der X-Achse

$$\left(\frac{1}{4}, \frac{3}{4}, \infty, \frac{1}{2}\right) \quad \text{und} \quad \left(\frac{1}{2}, \frac{3}{2}, \infty, 1\right)$$

harmonische Quadrupel bilden, womit die Verbindung zwischen den Konstruktionselementen der Skalen hergestellt ist (was übrigens auch mittels den Diagonallinien des Gitters hätte geschehen können). Durch Wiederholung der gleichen Betrachtung ergibt sich hieraus ohne Mühe, daß *irgend drei Skalenpunkte* $x_1\, x_2\, x_3$, für die

$$x_3 - x_2 = x_2 - x_1$$

ist, *mit* $x = \infty$ *harmonische Quadrupel*

$$(ABCD) = (x_1\, x_3\, \infty\, x_2)$$

bilden. Hieraus aber folgt unser Satz mit ausschließlicher Benutzung von Koordinatenlinien als Projektions-Strahlen zunächst für irgend drei aufeinanderfolgende und damit für irgend drei Gitterpunkte einer γ-Linie.

Bei unsrer Einführung von Koordinaten auf Grund von Vierseits-Konstruktionen erscheinen die Zahlen in Gestalt von *dyadischen „Systembrüchen“* (II, 2, mit der Basis $b = 2$). Beträgt die *Maschenweite* unsres Gitters $\frac{1}{2^k}$, so haben *alle* Koordinaten die Form[21]):

$$\sigma_k = g + \frac{a_1}{2} + \frac{a_2}{2^2} + \ldots + \frac{a_k}{2^k},$$

wo also g eine ganze Zahl und die a_ν 0 oder 1 bedeuten.

Da wir nun die Maschenweite unsres (offenen) Gitters durch fortgesetzte Vierseits-Konstruktionen unter jede Grenze herabdrücken können, so folgt aus unserm Satze über die Punkte eines (offenen) Gitters, daß *jede g-Linie durch eine lineare Gleichung in den Koordinaten x, y dargestellt werden kann, und daß umgekehrt jede derartige lineare Gleichung eine g-Linie darstellt.*

Damit aber ist der Beweis der Gültigkeit der projektiven Geometrie für ein beliebiges, den projektiven Axiomen I, 1—10 genügendes Kontinuum erbracht: die projektive Geometrie im begrenzten Gebiet oder, was dasselbe besagt, im offenen Kontinuum ist auf Grund der ihr eigentümlichen Axiome[22]) allein begründet, und damit auch der Fundamentalsatz erwiesen.

Darmstadt, den 6. April 1929.

---

[21]) Die von Hölder a. a. O., S. 205 und Klein-Rosemann, Vorlesungen über Nicht-Euklidische Geometrie, Berlin 1928, S. 160 verwandte Form $\frac{p}{2^n}$ ist für unsern Schluß durch *vollständige Induktion* ungeeignet.

[22]) Da zu diesen auch ein *Stetigkeitsaxiom* gehört (I, 7), so zählen wir es mit zu den „projektiven“ Axiomen (im Gegensatz zu anderen Autoren).

(Eingegangen am 10. 4. 1929.)

Über die topologische Erweiterung von Räumen.

Von

A. Tychonoff in Moskau.

Einleitung.

Die vorliegende Arbeit schließt an die abstrakt-topologischen Untersuchungen von Alexandroff und Urysohn an[1]) und behandelt das Problem der Einbettung von beliebigen topologischen Räumen in absolut abgeschlossene[2]) und insbesondere in bikompakte Räume. Ich beweise in erster Linie den folgenden

Satz I. *Zu jeder Kardinalzahl τ gibt es einen nur von dieser Kardinalzahl abhängenden bikompakten Raum R_τ von der Eigenschaft, daß jeder normale topologische Raum, der ein Umgebungssystem von einer Mächtigkeit $\leqq \tau$ besitzt, einer Teilmenge des Raumes R_τ homöomorph ist; dabei besitzt der Raum R_τ selbst ein Umgebungssystem von der Mächtigkeit τ. Wenn $\tau = \aleph_0$ ist, so ist R_τ dem Fundamentalquader des Hilbertschen Raumes homöomorph.*

(Man versteht bekanntlich unter dem Fundamentalquader des Hilbertschen Raumes die Gesamtheit aller Punkte $\xi = (x_1, x_2, \ldots, x_n, \ldots)$, deren Koordinaten x_n für jedes n den Ungleichungen $0 \leqq x_n \leqq \frac{1}{n}$ genügen.)

---

[1]) Siehe vor allem Alexandroff und Urysohn, Zur Theorie der topologischen Räume, Math. Annalen **92** (1924), S. 258, und Alexandroff, Über stetige Abbildungen kompakter Räume, Math. Annalen **96** (1926), S. 555; die Kenntnis dieser Arbeit (insbesondere auch die dort gebrauchte Terminologie) wird im folgenden vorausgesetzt; die erste dieser Arbeiten wird kurz durch „Alexandroff-Urysohn", die zweite durch „Alexandroff" zitiert. Wegen ausführlicher Darstellung der erwähnten Untersuchungen möge insbesondere auf „Mémoire sur les espaces compacts" (Verh. K. Akademie Amsterdam, Deel XIV, No. 1 (1929)) derselben Verfasser hingewiesen sein.

[2]) Alexandroff-Urysohn, S. 261.

In dem letzten Teile des soeben ausgesprochenen Satzes sind insbesondere die beiden bekannten Urysohnschen Metrisationssätze[3]) (der Metrisationssatz für separable normale Räume, sowie der Metrisationssatz für kompakte topologische Räume) und auch der Urysohnsche Satz von der Einbettbarkeit jedes separablen metrischen Raumes in den Fundamentalquader des Hilbertschen Raumes[4]) enthalten.

Die Frage liegt nahe, ob die Normalität eines topologischen Raumes auch notwendig ist, um den Raum in einen bikompakten topologischen Raum einbetten zu können. In der vorliegenden Arbeit (§ 4) wird auf diese Frage eine negative Antwort gegeben, indem gezeigt wird, daß auch nichtnormale topologische Räume als Teilmengen von bikompakten Räumen auftreten können. Somit entsteht die Frage, notwendige und hinreichende Bedingungen aufzustellen, damit ein Raum in einen bikompakten topologischen eingebettet werden kann. Um diese Bedingung formulieren zu können, führen wir die folgende Definition[5]) ein.

Ein topologischer Raum heißt *vollständig regulär*, wenn zu jedem Punkte x_0 und zu jeder ihn nicht enthaltenden abgeschlossenen Menge A eine im ganzen Raume stetige Funktion $f(x)$ definiert werden kann, die im ganzen Raume der Bedingung

$$0 \leqq f(x) \leqq 1$$

genügt und die überdies in x_0 gleich 0 und in sämtlichen Punkten von A gleich 1 ist.

Jeder vollständig reguläre Raum ist, wie leicht ersichtlich, regulär, während jeder normale Raum vollständig regulär ist. Im § 4 der vorliegenden Arbeit wird gezeigt, daß es vollständig reguläre Räume gibt, die nicht normal, und daß es reguläre Räume gibt, die nicht vollständig regulär sind, womit die Klasse der vollständig regulären Räume als eine logisch berechtigte Klasse von Räumen erscheint. Ihre mathematische Berechtigung wird durch den folgenden Satz gegeben:

Satz II. *Ein Raum ist dann und nur dann einer Teilmenge eines bikompakten topologischen Raumes homöomorph, wenn er vollständig regulär ist.*

Was endlich den allgemeinen Fall topologischer Räume betrifft, so gilt der

Satz III. *Jeder topologische Raum R ist einer Teilmenge eines absolut abgeschlossenen Raumes homöomorph, der ein Umgebungssystem besitzt,*

---

[3]) Siehe Urysohn, Zum Metrisationsproblem, Math. Annalen **94** (1925), S. 309, wo auch andere Arbeiten über denselben Gegenstand angegeben sind.

[4]) Siehe Urysohn, Der Hilbertsche Raum ..., Math. Annalen **92** (1924), S. 302.

[5]) Diese Definition rührt von Urysohn her. Vgl. seine Arbeit: Über die Mächtigkeit der zusammenhängenden Mengen, Math. Annalen **94** (1925), S. 292.

welches dieselbe Mächtigkeit hat wie ein gegebenes Umgebungssystem des Raumes R.

In der Formulierung der obigen Sätze wird unter einem Umgebungssystem selbstverständlich ein volles (d. h. dem System aller im Raume vorhandener Gebiete gleichwertiges) System von Umgebungen verstanden. Ein solches Umgebungssystem soll im folgenden kurz eine *Basis* genannt werden.

Ich möchte zum Schluß noch erwähnen, daß ich bei der endgültigen Redaktion dieser Arbeit von Herrn Alexandroff unterstützt worden bin.

§ 1.

Konstruktion der Räume R_τ.

Es sei $\{J_\alpha\}$ eine Menge von der Mächtigkeit τ von abstrakt gegebenen zueinander fremden Einheitsstrecken[6]) $0 \leqq t \leqq 1$. Ein Punkt x des Raumes R_τ ist definitionsgemäß der Inbegriff $\{t_1, t_2, \ldots, t_\alpha, \ldots\}$ von „Koordinaten" t_α, wobei t_α ein Punkt von J_α, also eine reelle Zahl $0 \leqq t_\alpha \leqq 1$ ist. Die Umgebungsdefinition geschieht folgendermaßen: Es sei $x_0 = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$ ein Punkt von R_τ. Wir wählen beliebig endlichviele $J_{\alpha_1}, J_{\alpha_2}, \ldots, J_{\alpha_k}$ und auf jedem dieser Intervalle J_{α_i} zwei rationale Zahlen $\tau'_{\alpha_i} < t^0_{\alpha_i} < \tau''_{\alpha_i}$; eine Umgebung von $x_0 = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$ besteht dann definitionsgemäß aus allen Punkten $x = \{t_1, t_2, \ldots, t_\alpha, \ldots\}$, die den Bedingungen $\tau'_{\alpha_i} < t_{\alpha_i} < \tau''_{\alpha_i}$ genügen.

Dieses Umgebungssystem hat, wie leicht ersichtlich, die Mächtigkeit $\aleph_0 \cdot \tau = \tau$. Ein gleichwertiges Umgebungssystem erhält man, wenn man für τ'_{α_i} bzw. τ''_{α_i} die Punkte $t^0_{\alpha_i} - \varepsilon_i$ bzw. $t^0_{\alpha_i} + \varepsilon_i$ wählt (ε_i sind beliebige hinreichend kleine positive Zahlen). Wir werden kurz schreiben $t_{\alpha_i} \subset S(t^0_{\alpha_i}; \varepsilon_i)$. Wenn $\varepsilon_i = \frac{1}{n_i}$ ist, werden wir die soeben definierte Umgebung durch

$$U^{n_1 n_2 \ldots n_k}_{\alpha_1 \alpha_2 \ldots \alpha_k}(x_0)$$

bezeichnen.

Der mittels dieses Umgebungssystems definierte Raum genügt, wie man leicht zeigt, allen Hausdorffschen Axiomen; R_τ ist also ein topologischer Raum.

Es ist nun leicht zu zeigen, daß $R_{\aleph_0}$ dem kompakten, durch die Bedingungen $0 \leqq x_n \leqq \frac{1}{n}$ bestimmten Fundamentalquadern des Hilbertschen Raumes homöomorph ist. Um diese Homöomorphie festzustellen, lassen wir jedem Punkte $x = \{t_1, \ldots, t_n, \ldots\}$ des Raumes $R_{\aleph_0}$ den Punkt

---

[6]) α nimmt alle Werte von Null bis zur ersten Ordnungszahl $\omega^{(\tau)}$ von der Mächtigkeit τ an.

$x^* = \left\{t_1, \ldots, \frac{t_n}{n}, \ldots\right\}$ des Hilbertschen Raumes entsprechen. Die Abbildung ist eineindeutig. Um zu zeigen, daß sie beiderseits stetig ist, muß man folgende Behauptungen beweisen:

a) Zu jeder Umgebung U von x_0 im $R_{\aleph_0}$ gibt es eine Zahl ε von der Eigenschaft, daß aus

$$\varrho(x_0^*; x^*) < \varepsilon$$

im Hilbertschen Raume die Inklusion

$$x \subset U(x_0) \qquad (\text{in } R_{\aleph_0})$$

folgt.

b) Zu jeder Zahl $\varepsilon > 0$ gibt es eine Umgebung $U(x_0)$ von der Eigenschaft, daß aus

$$x \subset U(x_0) \qquad (\text{in } R_{\aleph_0})$$

die Ungleichung

$$\varrho(x_0^*; x^*) < \varepsilon$$

im Hilbertschen Raum folgt.

Beweis von a). Es sei eine Umgebung $U(x_0)$ durch die Bedingungen

$$t_{n_i} \subset S(t_{n_i}^{(0)}; \varepsilon_i) \qquad (i = 1, 2, \ldots, k)$$

gegeben. Man wähle N größer als alle n_i, und ε kleiner als alle $\frac{\varepsilon_i}{N}$.

Wenn nun

$$\varrho(x_0^*; x^*) = \sqrt{\sum_{n=1}^{\infty} \frac{(t_n^0 - t_n)^2}{n^2}} < \varepsilon$$

ist, so ist auch

$$\frac{|t_n^{(0)} - t_n|}{n} < \varepsilon.$$

Es ist also für jedes n_i

$$|t_{n_i}^0 - t_{n_i}| < \varepsilon n_i < \varepsilon N < \varepsilon_i$$

und somit $x \subset U(x_0)$.

Beweis von b). Es sei $\varepsilon > 0$ gegeben; man wähle N so groß, daß

$$\sum_{n=N+1}^{\infty} \frac{1}{n^2} < \frac{\varepsilon^2}{2}$$

ist. Dann ist für jeden Punkt x der durch die Bedingungen

$$t_i \subset S\left(t_i^0; \frac{\varepsilon\sqrt{3}}{\pi}\right) \qquad (i = 1, 2, \ldots, N)$$

bestimmten Umgebung U von x_0

$$\varrho(x_0^*; x^*) = \sqrt{\sum_{n=1}^{\infty} \frac{(t_n^0 - t_n)^2}{n^2}} < \sqrt{\sum_{n=1}^{N} \frac{(t_n^0 - t_n)^2}{n^2} + \frac{\varepsilon^2}{2}} < \sqrt{\frac{\varepsilon^2}{2} \cdot \frac{6}{\pi^2} \sum_{n=1}^{N} \frac{1}{n^2} + \frac{\varepsilon^2}{2}} < \varepsilon,$$

w. z. b. w.

§ 2.
Beweis der Bikompaktheit von R_τ.

Vorbemerkung 1. Es sei irgendein Raum R und eine (abstrakt gedachte) Menge M (beliebiger Mächtigkeit) gegeben. Es sei ferner $\mathfrak{a}$ eine eindeutige (im allgemeinen nicht eineindeutige) Abbildung von M auf eine echte oder unechte Teilmenge $\mathfrak{a}(M)$ von R. Dann ist für jeden Punkt x von $\mathfrak{a}(M)$ eine Kardinalzahl, nämlich die Mächtigkeit der Menge aller Originalpunkte von x bei der Abbildung $\mathfrak{a}$, bestimmt. Diese Kardinalzahl nenne ich das *Gewicht* des Punktes x. Die Summe von Gewichten aller Punkte einer Teilmenge von $\mathfrak{a}(M)$ soll das Gewicht dieser Teilmenge heißen.

Wenn ξ ein Punkt von R ist, so nennen wir ihn *Konzentrationspunkt von M (in bezug auf* $\mathfrak{a}$), wenn, wie auch die Umgebung $U(\xi)$ von ξ gewählt sei, das Gewicht der Durchschnittsmenge $U(\xi)\cdot\mathfrak{a}(M)$ gleich der Mächtigkeit von M ist.

Die Menge aller Konzentrationspunkte von M in bezug auf $\mathfrak{a}$ ist eine in R abgeschlossene Menge, die wir durch $[M]_{\mathfrak{a}}$ bezeichnen werden.

Wenn R bikompakt ist, so besitzt jede unendliche Menge M in bezug auf jede Abbildung $\mathfrak{a}$ mindestens einen Konzentrationspunkt (Beweis mit Hilfe des Borel-Lebesgueschen Satzes ohne Schwierigkeit zu führen).

Im folgenden wird diese Bemerkung unter der Voraussetzung, daß R ein abgeschlossenes Intervall ist, benutzt.

Es sei E irgendeine unendliche Teilmenge von R_τ, deren Mächtigkeit $\mathfrak{m} \geqq \aleph_0$ sein möge. Indem man die Gesamtheit der ersten Koordinaten aller Punkte von E betrachtet, erhält man eine eindeutige Abbildung $\mathfrak{a}_1$ von E auf die Einheitsstrecke; wir bezeichnen durch t_1^0 einen (beliebig, aber eindeutig gewählten) Konzentrationspunkt von E in bezug auf $\mathfrak{a}_1$.

Es seien jetzt die reellen Zahlen

$$t_1^0,\ t_2^0,\ \ldots,\ t_\alpha^0,\ \ldots \qquad (0 \leqq t_\alpha \leqq 1)$$

für alle $\alpha < \alpha_0$ definiert (dabei ist α_0 eine beliebige Ordnungszahl unter ω^τ, und ω^τ die erste Ordnungszahl von der Mächtigkeit τ), und zwar so, daß folgende Bedingung erfüllt ist.

(A) Wenn E_n^α die Menge aller Punkte von E bedeutet, deren α-te Koordinate t_α der Ungleichung $|t_\alpha - t_\alpha^0| < \frac{1}{n}$ genügt, so hat der Durchschnitt eines beliebigen endlichen Systems

$$E_{n_1}^{\alpha_1}\cdot E_{n_2}^{\alpha_2}\cdot\ldots\cdot E_{n_k}^{\alpha_k} \qquad (\alpha_1 \leqq \alpha_2 \leqq \cdots \leqq \alpha_k < \alpha_0)$$

die Mächtigkeit $\mathfrak{m}$.

(Für $\alpha_0 = 2$ ist die Bedingung (A) erfüllt, weil dann

$$\alpha_1 = \alpha_2 = \ldots = \alpha_k = 1$$

ist und $E_{n_1}^{\alpha_1} \cdot E_{n_2}^{\alpha_2} \cdot \ldots \cdot E_{n_k}^{\alpha_k} = E_\nu^1$, wo ν die größte unter den Zahlen $n_1, n_2, \ldots, n_k$ ist; die Menge E_ν^1 hat offenbar die Mächtigkeit $\mathfrak{m}$, weil t_1^0 ein Konzentrationspunkt von E ist.)

Es sei jetzt $\mathfrak{a}$ die Abbildung von R_τ auf die Einheitsstrecke, die dadurch entsteht, daß man jedem Punkte von R_τ seine α_0-te Koordinate entsprechen läßt.

Man betrachte alle abgeschlossenen Mengen von der Gestalt

$$[E_{n_1}^{\alpha_1} \cdot E_{n_2}^{\alpha_2} \cdot \ldots \cdot E_{n_k}^{\alpha_k}]_{\mathfrak{a}} = F_{n_1 n_2 \ldots n_k}^{\alpha_1 \alpha_2 \ldots \alpha_k},$$

wobei wie immer $\alpha_1 \leqq \alpha_2 \leqq \ldots \leqq \alpha_k < \alpha_0$ ist.

Ich behaupte, daß je endlichviele unter den Mengen $F_{n_1 n_2 \ldots n_k}^{\alpha_1 \alpha_2 \ldots \alpha_k}$ einen nicht leeren Durchschnitt haben. Es seien in der Tat

$$F_{n_1^1 n_2^1 \ldots n_{k_1}^1}^{\alpha_1^1 \alpha_2^1 \ldots \alpha_{k_1}^1}, \quad F_{n_1^2 n_2^2 \ldots n_{k_2}^2}^{\alpha_1^2 \alpha_2^2 \ldots \alpha_{k_2}^2}, \quad \ldots, \quad F_{n_1^h n_2^h \ldots n_{k_h}^h}^{\alpha_1^h \alpha_2^h \ldots \alpha_{k_h}^h}$$

beliebig gegeben. Dann ist

$$(1) \qquad \prod_{i=1}^{h} F_{n_1^i n_2^i \ldots n_{k_i}^i}^{\alpha_1^i \alpha_2^i \ldots \alpha_{k_i}^i} = \prod_{i=1}^{h} \left[E_{n_1^i n_2^i \ldots n_{k_i}^i}^{\alpha_1^i \alpha_2^i \ldots \alpha_{k_i}^i} \right]_{\mathfrak{a}}.$$

Nun ist aber (wie die Mengen A und B gewählt sein mögen) stets

$$[A]_{\mathfrak{a}} \cdot [B]_{\mathfrak{a}} \supset [A \cdot B]_{\mathfrak{a}}.$$

Aus dieser Bemerkung und der Identität (1) folgt somit

$$(2) \qquad \prod_{i=1}^{h} F_{n_1^i n_2^i \ldots n_{k_i}^i}^{\alpha_1^i \alpha_2^i \ldots \alpha_{k_i}^i} \supset \left[\prod_{i=1}^{h} \prod_{j=1}^{k_i} E_{n_j^i}^{\alpha_j^i} \right]_{\mathfrak{a}}.$$

Es gilt ferner vermöge der Bedingung (A):

$$(3) \qquad \text{Mächt. von} \left[\prod_{i=1}^{h} \prod_{j=1}^{k_i} E_{n_j^i}^{\alpha_j^i} \right] = \mathfrak{m},$$

woraus folgt, daß

$$\left[\prod_{i=1}^{h} \prod_{j=1}^{k_i} E_{n_j^i}^{\alpha_j^i} \right]_{\mathfrak{a}}$$

und insbesondere

$$\prod_{i=1}^{h} F_{n_1^i n_2^i \ldots n_{k_i}^i}^{\alpha_1^i \alpha_2^i \ldots \alpha_{k_i}^i}$$

nicht leer ist.

Da je endlichviele unter den Mengen $F_{n_1 n_2 \dots n_k}^{\alpha_1 \alpha_2 \dots \alpha_k}$ $(\alpha_1 \leqq \alpha_2 \leqq \dots \leqq \alpha_k < \alpha_0)$ einen nicht leeren Durchschnitt haben, so haben alle diese Mengen (die ja abgeschlossene Teilmengen der Strecke $0 \leqq t \leqq 1$ sind) mindestens einen Punkt gemein. Einen beliebig (aber eindeutig) gewählten unter den Punkten der Durchschnittsmenge aller $F_{n_1 n_2 \dots n_k}^{\alpha_1 \alpha_2 \dots \alpha_k}$ bezeichnen wir durch $t_{\alpha_i}^0$, und man zeigt leicht, daß die Bedingung (A) für alle $\alpha_1 \leqq \alpha_2 \leqq \dots \leqq \alpha_k < \alpha_0 + 1$ erfüllt ist [7]).

Auf diese Weise bestimmt man für jede Ordnungszahl $\alpha < \omega^\tau$ einen Punkt t_α^0 der Strecke $0 \leqq t \leqq 1$. Man betrachte jetzt den Punkt x_0 von R_τ mit den Koordinaten

$$t_1^0, t_2^0, \dots, t_\alpha^0, \dots \qquad (\alpha < \omega^\tau):$$

$$x_0 = \{t_1^0, t_2^0, \dots, t_\alpha^0, \dots\}.$$

Ich behaupte, daß x_0 ein vollständiger Häufungspunkt der Menge E ist. In der Tat enthält eine beliebige Umgebung von x_0 $U_{\alpha_1 \alpha_2 \dots \alpha_k}^{n_1 n_2 \dots n_k}(x_0)$ die Menge $E_{n_1}^{\alpha_1} \cdot E_{n_2}^{\alpha_2} \cdot \ldots \cdot E_{n_k}^{\alpha_k}$, die (vermöge der Bedingung (A)) die Mächtigkeit $\mathfrak{m}$ hat, w. z. b. w.

§ 3.
Beweis der Sätze I und II.

Um die Sätze I und II zu beweisen, genügt es zu zeigen:

1. daß jede Teilmenge eines bikompakten Raumes vollständig regulär ist;

2. daß jeder vollständig reguläre Raum R mit einer Basis von der Mächtigkeit τ einer Teilmenge von R_τ homöomorph ist.

Die erste Behauptung folgt daraus, daß die vollständige Regularität eine im Urysohnschen Sinne transitive Eigenschaft ist, d. h. daß eine Teilmenge eines vollständig regulären Raumes, als Raum betrachtet, selbst vollständig regulär ist. Es sei in der Tat R eine Teilmenge eines vollständig regulären Raumes R^*. Ist nun x_0 ein Punkt und A eine zu x_0 fremde in R abgeschlossene Menge, so gibt es eine in R^* abgeschlossene Menge A^* derart, daß $A = R \cdot A^*$. Da R^* vollständig regulär ist, kann man dort eine stetige Funktion $f(x)$ erklären, die folgenden Bedingungen genügt:

$$\begin{cases} \quad f(x_0) = 0, & \\ \quad f(x) = 1 & \text{für alle } x \subset A^*, \\ 0 \leqq f(x) \leqq 1 & \text{für alle übrigen Punkte.} \end{cases}$$

---

[7]) Es seien in der Tat

$$E_{n_1}^{a_1}, E_{n_2}^{a_2}, \dots, E_{n_k}^{a_k}$$

gegeben. $E^* = E_{n_1}^{a_1} \cdot E_{n_2}^{a_2} \cdot \ldots \cdot E_{n_k}^{a_k}$ hat die Mächtigkeit $\mathfrak{m}$, und $t_{a_0}^{(0)}$ ist ein Konzentrationspunkt dieser Menge in bezug auf $\mathfrak{a}$; daraus und aus der Definition von $E_{n_0}^{a_0}$ folgt dann unmittelbar, daß bei jedem n_0 $E^* \cdot E_{n_0}^{a_0}$ von derselben Mächtigkeit $\mathfrak{m}$ ist.

Die Menge A ist in der Menge A^* enthalten, also ist für alle ihre Punkte $f(x)=1$. Betrachtet man die Funktion $f(x)$ nur in Punkten von R, so genügt sie allen aufgestellten Bedingungen, womit die vollständige Regularität von R bewiesen ist.

Bemerkung 2. Ist R ein vollständig regulärer Raum, so gibt es für jeden Punkt x_0 und jede seine Umgebung $U(x_0)$ eine andere Umgebung desselben Punktes $V(x_0)$ derart, daß eine stetige Funktion $f(x)$ sich erklären läßt, so daß

$$\begin{cases} f(x)=0, & \text{wenn } x \subset \overline{V(x_0)}, \\ f(x)=1, & \text{wenn } x \subset R-U(x_0), \\ 0 \leqq f(x) \leqq 1 & \text{für alle übrigen Punkte von } R. \end{cases}$$

Um letztere Tatsache zu beweisen, erkläre man im Raume R eine stetige Funktion $f^*(x)$ dadurch, daß man

$$\begin{cases} f^*(x_0)=0, & \\ f^*(x)\;=1, & \text{wenn } x \subset R-U(x), \\ 0 \leqq f^*(x) \leqq 1 & \text{in allen übrigen Punkten von } R \end{cases}$$

setzt.

Die Funktion $f(x)$ ist stetig, es gibt also eine Umgebung von x_0, $V(x_0)$, für deren Punkte $f^*(x) < \frac{1}{2}$. Dann kann man eine Funktion $f(x)$ z. B. folgendermaßen definieren:

$$\begin{cases} f(x)=f^*(x) & , \text{ falls } f^*(x)=1, \\ f(x)=0 & , \text{ falls } f^*(x) \leqq \frac{1}{2}, \\ f(x)=2\left(f^*(x)-\frac{1}{2}\right), & \text{ falls } \frac{1}{2} < f^*(x) < 1. \end{cases}$$

Der Beweis der Behauptung 2 geschieht nunmehr mit Hilfe einer von Urysohn zum Beweis eines seiner Metrisationssätze angewandten Methode. Es sei R ein vollständig regulärer Raum mit einer Basis $\mathfrak{B}$ von der Mächtigkeit τ. Ein Paar zur Basis $\mathfrak{B}$ gehörender Umgebungen $(U, V)=\pi$ soll kanonisch heißen, falls $U \supset \overline{V}$ ist und eine stetige Funktion $f(x)$ im Raume R bestimmt werden kann, die in $\overline{V}$ verschwindet, in $R-U$ gleich 1 ist und überall in R der Ungleichung $0 \leqq f(x) \leqq 1$ genügt.

Es seien

$$\pi_1, \pi_2, \ldots, \pi_\alpha, \ldots \qquad (\alpha < \omega^\tau)$$

alle kanonischen Paare[8]) und

[8]) Die Mächtigkeit der Menge aller kanonischen Paare ist höchstens gleich der Mächtigkeit der Basis; sie kann aber auch nicht kleiner sein, weil (vermöge der Bemerkung 2) jedes U aus $\mathfrak{B}$ mindestens in einem π vorkommt; die Mächtigkeit aller kanonischen Paare ist also genau τ.

$$f_1, f_2, \ldots, f_\alpha, \ldots \qquad (\alpha < \omega^\tau)$$

die ihnen entsprechenden stetigen Funktionen.

Wir lassen nun einem Punkt $x \subset R$ den Punkt

$$f(x) = x^* = \{t_1, t_2, \ldots, t_\alpha, \ldots\}$$

entsprechen, wo $t_\alpha = f_\alpha(x)$. Die dadurch bestimmte Abbildung von R auf eine Teilmenge R^* von R_τ ist eineindeutig. Sind in der Tat x und y verschiedene Punkte von R, so gibt es zunächst eine Umgebung U, die x, aber nicht y enthält; vermöge der Bemerkung 2 gibt es sodann eine in U enthaltene Umgebung V von x, so daß U und V ein kanonisches Paar $\pi_\alpha = (U, V)$ bilden. Somit ist $f_\alpha(x) = 0$, während $f_\alpha(y) = 1$ ist; die Punkte x^* und y^* sind also verschieden.

Die Abbildung $x^* = f(x)$ ist *beiderseits* stetig.

1. Es sei x_0 ein Punkt,

$$f(x_0) = x_0^* = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$$

sein Bildpunkt in R^*, $U(x_0^*) = U_{\alpha_1 \alpha_2 \ldots \alpha_k}^{n_1 n_2 \ldots n_k}$ eine beliebige Umgebung letzteren Punktes. Um die Stetigkeit der Funktion $x^* = f(x)$ nachzuweisen, genügt es eine den Punkt x_0 enthaltende offene Menge G zu finden, dessen Bild $f(G)$ in $U(x_0^*)$ enthalten ist. Man beachte zu diesem Zwecke die Menge $G_{\alpha_i}^{n_i}$ aller Punkte x von R, für die

$$(1_i) \qquad t_{\alpha_i}^0 - \frac{1}{n_i} < f_{\alpha_i}(x) < t_{\alpha_i}^0 + \frac{1}{n_i} \qquad (1 \leq i \leq k)$$

ist; da $f_\alpha(x)$ stetig ist, so ist G_α^n eine offene Menge; da außerdem $f_{\alpha_i}(x_0) = t_{\alpha_i}^0$ ist, so ist $x_0 \subset G_{\alpha_i}^{n_i}$. Es genügt jetzt

$$G = G_{\alpha_1}^{n_1} \cdot G_{\alpha_2}^{n_2} \cdot \ldots \cdot G_{\alpha_k}^{n_k}$$

zu setzen. G ist eine offene Menge und für $x \subset G$ sind alle Bedingungen (1_i) erfüllt, d. h. es ist $f(x) \subset U(x_0^*)$; da überdies $x_0 \subset G$ ist, ist unsere Behauptung bewiesen.

2. Um die Stetigkeit der Umkehrungsfunktion $x = f^{-1}(x^*)$ zu beweisen, zeigen wir, daß für jede $U(x_0)$ sich eine $U(x_0^*)$ finden läßt, für welche $f^{-1}(U(x_0^*)) \subset U(x_0)$. Letztere Bedingung heißt aber nichts anderes, als daß aus

$$y \subset R - U(x_0)$$
$$y^* \subset R^* - U(x_0^*)$$

folgen soll. Es ist nun $U(x_0)$ beliebig gegeben; man bestimme ein V so, daß $x_0 \subset V$ und (U, V) ein kanonisches Paar, und zwar $\pi_\alpha = (U, V)$ ist; dann ist aber $f_\alpha(x_0) = 0$; $f_\alpha(y) = 1$ für jedes $y \subset R - U$, d. h. die α-te Koordinate von x_0^* ist gleich Null, die α-te Koordinate eines jeden Punktes

y^* für $y \subset R - U$ ist dagegen gleich 1. Wenn jetzt $U(x_0^*) = U_\alpha^2$ gesetzt wird, so folgt aus $y \subset R - U(x_0)$, daß die α-te Koordinate von y^* gleich 1; also $> \frac{1}{2}$ ist; y^* liegt somit außerhalb von $U(x_0^*)$.

Die Homöomorphie der beiden Räume R und R^* ist hiermit nachgewiesen. Dadurch sind aber auch die beiden Sätze I und II bewiesen.

§ 4.

Konstruktion von Beispielen.

In diesem Paragraphen will ich zeigen:

1. daß es vollständig reguläre Räume gibt, die nicht normal sind, und
2. daß es reguläre Räume gibt, die nicht vollständig regulär sind.

Ich konstruiere zunächst einen Raum S folgendermaßen. Die Punkte x von S sind Paare (α, β), wobei α und β Ordnungszahlen sind, und zwar durchläuft α alle Werte $\leqq \omega_1$ und β alle Werte $\leqq \omega$. Wenn $x_0 = (\alpha_0, \beta_0)$ $(1 \leqq \alpha_0 \leqq \omega_1,\ 1 \leqq \beta_0 \leqq \omega)$ ein Punkt von S ist, so erhält man eine beliebige Umgebung $U_{\alpha'\beta'}(x_0)$ $0 \leqq \alpha' < \alpha_0$, $0 \leqq \beta < \beta_0$ dadurch, daß man alle $x = (\alpha; \beta)$ betrachtet, für die $\alpha' < \alpha \leqq \alpha_0$, $\beta' < \beta \leqq \beta_0$ ist.

Wie leicht ersichtlich, ist S ein bikompakter Raum.

Man bezeichne nun durch $\dot{S}$ den Raum, der aus S durch Tilgen des Punktes (ω_1, ω) entsteht. Als Teilmenge eines bikompakten Raumes S ist $\dot{S}$ vollständig regulär; andererseits ist aber $\dot{S}$ nicht normal, weil die beiden abgeschlossenen und zueinander fremden Teilmengen von $\dot{S}$

$$X = \left\{ \begin{matrix} (\alpha, \omega) \\ 1 \leqq \alpha < \omega_1 \end{matrix} \right\} \quad \text{und} \quad Y = \left\{ \begin{matrix} (\omega_1, \beta) \\ 1 \leqq \beta < \omega \end{matrix} \right\}$$

durch keine zwei zueinander fremde Gebiete getrennt werden können. Es gilt sogar die schärfere

Behauptung I. Wenn G ein die Menge $Y_n = \left\{ \begin{matrix} (\omega_1, \beta) \\ n \leqq \beta < \omega \end{matrix} \right\}$ enthaltendes Gebiet ist, so gibt es ein $\alpha_0 < \omega_1$ derart, daß $\overline{G}$ alle Punkte $X_{\alpha_0} = \{\alpha, \omega\}$, $\alpha_0 \leqq \alpha < \omega_1$, enthält.

In der Tat gibt es für jeden Punkt $x = (\omega_1, \beta)$ eine $U_{\alpha_\beta, \beta'}(x) \subset G$; also sind insbesondere alle $x = (\alpha, \beta)$ $\alpha > \alpha_\beta$ in G enthalten. Da die Menge aller α_β abzählbar ist, gibt es eine Zahl α_0, die kleiner als ω_1 und größer als alle α_β ist. Somit sind alle Punkte (α, β), $\alpha > \alpha_0$, $n \leqq \beta < \omega$ in G und alle Punkte (α, ω), $\alpha > \alpha_0$ in $\overline{G}$ enthalten, w. z. b. w.

Ein regulärer Raum ist (vermöge des Satzes II) dann und nur dann vollständig regulär, wenn er in einen bikompakten Raum eingebettet werden kann. Da jeder bikompakte Raum regulär ist, so folgt aus dem soeben

Gesagten, daß ein nicht bikompakter regulärer Raum, der in bezug auf reguläre Punkte abgeschlossen ist, sicher nicht vollständig regulär ist. Bevor wir einen solchen Raum konstruieren, machen wir die folgende

Behauptung II. Wenn $\dot{G}$ ein die Mengen

$$X_{\alpha_0} = \left\{ \begin{matrix} (\alpha, \omega) \\ \alpha_0 < \alpha < \omega_1 \end{matrix} \right\}, \quad Y_{n_0} = \left\{ \begin{matrix} (\omega_1, \beta) \\ n_0 < \beta < \omega \end{matrix} \right\}$$

enthaltendes Gebiet in $\dot{S}$ ist, so ist $\dot{S} - \dot{G}$ (als Relativraum in $\dot{S}$ betrachtet) bikompakt.

In der Tat ist

$$\dot{S} - \dot{G} = S - G$$

(wobei $G = \dot{G} + (\omega_1, \omega)$ ist), woraus unsere Behauptung folgt.

Es sei jetzt R^* ein topologischer Raum, der dadurch entsteht, daß man die Vereinigungsmenge von abzählbar vielen Exemplaren

$$\dot{S}^1, \dot{S}^2, \ldots, \dot{S}^k, \ldots$$

von $\dot{S}$ betrachtet. Wir bezeichnen die Punkte von $\dot{S}^k$ durch $(\alpha, \beta)^k$, die Mengen X_α und Y_n von $\dot{S}^k$ durch X_α^k bzw. Y_n^k usw. Man betrachte ferner die stetige Zerlegung[9]) von R^* in (höchstens aus zwei Punkten bestehende) Mengen X, wobei

X entweder ein Punkt $(\alpha, \beta)^k$, $\alpha \neq \omega_1$, $\beta \neq \omega$ von R

oder ein Punktepaar $[(\omega_1, \beta)^{2n-1}, (\omega_1, \beta)^{2n}]$, $n = 1, 2, \ldots$

oder ein Punktepaar $[(\alpha, \omega)^{2n}, (\alpha, \omega)^{2n+1}]$, $n = 1, 2, \ldots$

ist. Der durch diese Zerlegung induzierte Raum ist, wie leicht ersichtlich, ein regulärer Raum $\dot{R}^*$[10]). Jedem Punkte $(\alpha, \beta)^k$ von R^* entspricht ein Punkt $(\alpha, \beta)^{*k}$ von $\dot{R}^*$, ebenso entsprechen die Teilmengen $\dot{S}^k$ bzw. X_α^k bzw. Y_n^k von R^* Teilmengen $\dot{S}^{*k}$ bzw. X_α^{*k} bzw. Y_n^{*k} von $\dot{R}^*$.

Wir konstruieren endlich einen Raum

$$R = \dot{R}^* + \xi$$

dadurch, daß wir einen neuen Punkt ξ mittels der Umgebungserklärung

$$U_m(\xi) = \xi + \left(\dot{R}^* - \sum_{k=1}^{m} \dot{S}^{*k}\right) \qquad (m = 1, 2, \ldots)$$

einführen. Man zeigt ohne Schwierigkeit, daß R ein regulärer Raum ist.

Ich behaupte, daß R in bezug auf alle regulären Punkte abgeschlossen ist.

---

[9]) Alexandroff, Math. Annalen 96 (1926), S. 557 (§ 4).

[10]) Man könnte sagen, daß $\dot{R}^*$ aus R^* dadurch entsteht, daß man in R^* die Punkte $(\omega_1, \beta)^{2n-1}$ bzw. $(\alpha, \omega)^{2n}$ mit $(\omega_1, \beta)^{2n}$ bzw. $(\alpha, \omega)^{2n+1}$ identifiziert.

Um dies zu beweisen machen wir die folgende

Behauptung III. Es sei G ein Teilgebiet von S, welches die Menge X_{α_0} enthält; es gibt dann eine natürliche Zahl n_0 von der Eigenschaft, daß $\bar{G}$ alle Punkte (ω_1, β), $n_0 \leqq \beta < \omega$ (d. h. die Menge Y_{n_0} enthält.

Es existiert in der Tat für jedes $x = (\alpha, \omega)$, $\alpha > \alpha_0$, eine $U_{\alpha', n_\alpha}(x) \subset G$; infolgedessen sind alle (α, β), $n_\alpha < \beta < \omega$, in G enthalten; da es nur abzählbar viele verschiedene n_α gibt, gibt es ein $n_0 < \omega$, so daß für unabzählbar viele α alle (α, β), $n_0 < \beta < \omega$, in G enthalten sind; dann sind aber alle (ω_1, β), $n_0 < \beta < \omega$, in $\bar{G}$ enthalten.

Es sei jetzt

$$R + \eta$$

ein regulärer Raum. Dann gibt es zwei Umgebungen $U(\xi) = U_N(\xi)$ und $V(\eta)$, so daß

$$\overline{U(\xi)} \cdot \overline{V(\eta)} = 0$$

ist; wir dürfen dabei stets voraussetzen, daß N gerade ist. Man konstruiere ferner

$$\overline{V_1(\eta)} \subset V(\eta), \ldots, \overline{V_{m+1}(\eta)} \subset V_m(\eta), \ldots \quad (m = 1, 2, \ldots)$$

und setze

$$G_m = R - \overline{V_m(\eta)}.$$

Dann ist

$$\bar{G}_m \subset R - V_m(\eta) \subset R - \overline{V_{m+1}(\eta)} \subset G_{m+1}. \tag{0}$$

Außerdem ist

$$G_1 \supset U_N(\xi) \supset \sum_{k=N+1}^{\infty} S^{*k}, \tag{1}$$

so daß

$$V(\eta) \subset \sum_{k=1}^{N} S^{*k} + \eta$$

ist.

Aus (1) folgt weiter, daß

$$G_1 \supset X^{*N+1},$$

also auf Grund der Behauptung III existiert ein n_1, so daß

$$\bar{G}_1 \supset Y_{n_1}^{*N+1} = Y_{n_1}^{*N}$$

und also (vermöge (0))

$$G_2 \supset Y_{n_1}^{*N}.$$

Aus der Behauptung I folgt sodann, daß es ein α_2 gibt, so daß

$$G_3 \supset \bar{G}_2 \supset X_{\alpha_2}^{*N} = X_{\alpha_2}^{*N-1}.$$

Daraus folgt weiter (nach Behauptung III), daß

$$G_4 \supset \bar{G}_3 \supset Y_{n_3}^{*N-1} = Y_{n_3}^{*N-2},$$

was (nach Behauptung I)

$$G_5 \supset \overline{G}_4 \supset X^{*N-2}_{a_5}$$

ergibt.

Eine wiederholte Anwendung der beiden Behauptnngen I und III schließt mit

$$G_{N+2} \supset Y^{*1}_{a_N} + X^{*1}_{n_{N+1}}.$$

Nun ist

$$V_{N+2} \subset R - G_{N+2}.$$

Da aber G_{N+2} sämtliche $X^{*k}_{a_{N-k}}$ und $Y^{*k}_{n_{N-k}}$ $k \leq N$ und sämtliche $\dot{S}^{*m}$ $m > N+1$ enthält, so ist $V_{N+2}(\eta)$ in einer nach Bemerkung II bikompakten Teilmenge von R^* enthalten, was aber nur dann möglich[11]) ist, wenn η in V_{N+2} und somit in $R + \eta$ isoliert ist; damit ist die Abgeschlossenheit von R^* in bezug auf reguläre Punkte bewiesen.

Es sei hierzu noch bemerkt, daß durch das soeben konstruierte Beispiel eine von Alexandroff und Urysohn (a. a. O. S. 266) gestellte Frage, ob jeder reguläre nicht absolut abgeschlossene Raum sich zu einem ebenfalls regulären Raum durch Hinzufügung eines nicht isolierten Punktes erweitern läßt, eine volle (und zwar negative) Antwort erhalten hat.

§ 5.

Beweis des Satzes III.

Der Beweis des Satzes III beruht auf der Betrachtung von Zerlegungen[12]) der Räume R_τ in zueinander fremde abgeschlossene Mengen: es wird sich nämlich zeigen, daß die absolut abgeschlossenen Räume, deren Existenz im Satze III behauptet wird, durch Zerlegungen der Räume R_τ bestimmt werden. Bei der Untersuchung dieser Zerlegungen werden wir folgende von Alexandroff a. a. O. bewiesene Tatsache benutzen: ein durch eine Zerlegung eines bikompakten Raumes R bestimmter Raum R^* ist ein topologischer Raum; wenn R eine Basis von einer Mächtigkeit $\leq \mathfrak{m}$ besitzt, so gilt dasselbe auch für R^*. [13])

Es sei R ein topologischer Raum mit einer Basis $\mathfrak{B}$ von der Mächtigkeit $\leq \tau$. Der Satz III wird bewiesen sein, wenn wir gezeigt haben werden, daß es eine Zerlegung von R_τ gibt, die einen absolut abgeschlossenen

---

[11]) Alexandroff und Urysohn, S. 263, Satz V.

[12]) Im Sinne von Alexandroff, Math. Annalen 96 (1926), § 3 (S. 556—557). Es sei hier ausdrücklich erwähnt, daß wir (der Alexandroffschen Definition entsprechend) beliebige allgemeine und nicht notwendig etwa stetige Zerlegungen betrachten.

[13]) Bei Alexandroff ist diese Behauptung explizite nur für stetige Zerlegungen und $\mathfrak{m} = \aleph_0$ (S. 262, Satz V und § 13) bewiesen. Der Beweis gilt aber wörtlich auch in unserem allgemeinen Falle.

Raum bestimmt, welcher eine dem Raume R homöomorphe Teilmenge enthält. Wir gehen jetzt zum Beweise letzterer Behauptung über.

Es sei X_α eine auf der Einheitsstrecke liegende abgeschlossene Menge. Im folgenden werden wir die Mengen $X = \{X_1, X_2, \ldots, X_\alpha, \ldots\}$ des Raumes R_τ betrachten, die aus allen Punkten $x = \{t_1, t_2, \ldots, t_\alpha, \ldots\}$ bestehen, für die $t_\alpha \subset X_\alpha$ ist. Die Mengen X sind abgeschlossen. In der Tat, wenn $x_0 = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$ nicht zu X gehört, so ist wenigstens für ein α

$$\varrho(t_\alpha^0, X_\alpha) = \delta > 0$$

(die Entfernung auf J_α gemessen). Dann enthält aber seine, durch die Bedingung $t_\alpha \subset S\left(t_\alpha^0, \frac{\delta}{2}\right)$ bestimmte Umgebung keinen Punkt der Menge X; x_0 ist somit kein Häufungspunkt von X.

Jetzt wollen wir diejenigen Mengen X, die als Zerlegungseinheiten von R_τ auftreten sollen, definieren. Zu diesem Zwecke bestimmen wir für ein beliebiges Gebiet G des Raumes R zwei Funktionen $\varphi(x)$ und $\psi(x)$ durch die Vorschrift

$$\varphi_G(X) = \begin{cases} 0 & \text{falls } x \subset \overline{G} \\ 1 & \text{falls } x \subset G; \end{cases} \qquad \psi_G(X) = \begin{cases} 0 & \text{falls } x \subset G \\ 1 & \text{falls } x \subset R - G. \end{cases}$$

Für jeden Punkt des Raumes ist $\varphi_G(x) \leqq \psi_G(x)$. Man sieht auch leicht, daß, wenn $\varphi_G(x_0) = \psi_G(x)$ ist, eine Umgebung $U(x_0)$ des Punktes x_0 existiert, für deren jeden Punkt $y \subset U(x_0)$

$$\varphi_G(y) = \psi_G(y) = \varphi_G(x_0) = \psi_G(x_0)$$

ist.

Es seien

$$U_1, U_2, \ldots, U_\alpha, \ldots \qquad (\alpha < \omega^{(\tau)})$$

die Elemente der Basis $\mathfrak{B}$ des Raumes R. Einfachheitshalber setzen wir

$$\varphi_\alpha(x) = \varphi_{u_\alpha}(x) \quad \text{und} \quad \psi_\alpha(x) = \psi_{u_\alpha}(x).$$

Es sei ferner $X_\alpha(x)$ die Gesamtheit aller der Bedingung $\varphi_\alpha(x) \leqq t \leqq \psi_\alpha(x)$ genügender reeller Zahlen und

$$X(x) = \{X_1(x), X_2(x), \ldots, X_\alpha(x), \ldots\}.$$

Wenn x und y verschiedene Punkte von R sind, so gibt es eine U_α so, daß $x \subset U_\alpha$, $y \subset R - \overline{U}_\alpha$, woraus folgt, daß $X(x)$ und $X(y)$ zueinander fremd sind.

Wenn die Vereinigungsmenge Ξ der (für alle Punkte x von R) bestimmten $X(x)$ nicht mit dem ganzen Raum R_τ identisch ist, betrachten wir alle Punkte von $R - \Xi$ als neue Zerlegungseinheiten, so daß schließlich der ganze Raum R_τ in zueinander fremde abgeschlossene Mengen zerlegt wird: $R_\tau = \sum X$. Diese Zerlegung ist durch den Raum R allein ein-

deutig bestimmt und soll deshalb als die Zerlegung $\zeta(R_\tau, R)$ bezeichnet werden. Der durch $\zeta(R_\tau, R)$ bestimmte Raum soll R_τ^R heißen. Die Zerlegung $\zeta(R_\tau, R)$ induziert aber auch eine, im allgemeinen unstetige Abbildung von R_τ auf R_τ^R, die wir durch φ bezeichnen werden: $R_\tau^R = \varphi(R_\tau)$. Die Punkte von R_τ^R werden wir durch x^* bezeichnen. Da jedem Punkte $x < R$ eine Menge X, d. h. ein Punkt $x^* = f(x)$ zugeordnet ist, sind wir im Besitze einer eineindeutigen Abbildung $x^* = f(x)$ des Raumes R auf eine (echte oder unechte) Teilmenge R^* von R_τ^R. Wir werden zeigen, daß diese Abbildung eine Homöomorphie ist. Dazu brauchen wir aber eine bequeme Basis des Raumes R_τ^R.

Es sei $G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}(x_0)$ die Gesamtheit aller Punkte $x = \{t_1, t_2, \ldots, t_\alpha, \ldots\}$ von R_τ, für die $t_{\alpha_i} < S\left(X_{\alpha_i}(x_0), \frac{1}{n_i}\right)$ ist. $G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}$ ist ein Gebiet. Es sei ferner $V_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}(x_0^*)$ die Gesamtheit aller x^*, die, als Mengen X betrachtet, in $G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}$ liegen. Ich behaupte, daß die $V_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}$ eine Basis bilden. Um es einzusehen, genügt es zu zeigen, daß man für jedes Gebiet $G > X_0$ ein Gebiet $G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}(x_0)$ finden kann von der Eigenschaft, daß

$$(1) \qquad X_0 < G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k} < G$$

ist.

Um dieses $G_{\alpha_1\alpha_2\ldots\alpha_k}^{n_1n_2\ldots n_k}$ zu konstruieren, wähle man für jeden Punkt $x < X_0$ eine $U_{\alpha_1^x, \alpha_2^x, \ldots \alpha_k^x}^{2m_1^x, 2m_2^x, \ldots, 2m_k^x}(x)$, wobei die $\alpha_1^x, \alpha_2^x, \ldots, \alpha_k^x, m_1^x, m_2^x, \ldots, m_k^x$ durch die Bedingung

$$U_{\alpha_1^x, \alpha_2^x, \ldots, \alpha_k^x}^{m_1^x, m_2^x, \ldots, m_k^x}(x) < G$$

bestimmt sind. Auf diese Weise entsteht eine Überdeckung der abgeschlossenen Teilmenge X_0 von R_τ; nach dem Borel-Lebesgueschen Satze kann man diese Überdeckung durch eine endliche Teilüberdeckung ersetzen. Es seien $\alpha_1, \alpha_2, \ldots, \alpha_k$ alle bei den Elementen dieser Überdeckung auftretenden untere Indizes und $2m_i$ der größte unter den zu α_i gehörenden oberen Indizes. Dann ist

$$G_{\alpha_1, \alpha_2, \ldots, \alpha_k}^{2m_1, 2m_2, \ldots, 2m_k} < G.$$

Es sei in der Tat $\xi = \{t_1, t_2, \ldots, t_\alpha, \ldots\}$ irgendein Punkt von $G_{\alpha_1, \alpha_2, \ldots \alpha_k}^{2m_1, 2m_2, \ldots, 2m_k}$. Es ist $|t_{\alpha_i} - t_{\alpha_i}^0| < \frac{1}{2m_i}$, wobei

$$x_0 = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$$

ein Punkt von X_0 ist; nun ist aber x_0 in einer bestimmten $U_{\alpha_1^x, \alpha_2^x, \ldots, \alpha_k^x}^{2m_1^x, 2m_2^x, \ldots, 2m_k^x}(x)$ mit $x = \{t_1^x, t_2^x, \ldots, t_\alpha^x, \ldots\}$ enthalten. Alle $\alpha_1^x, \alpha_2^x, \ldots, \alpha_k^x$ kommen unter den $\alpha_1, \alpha_2, \ldots, \alpha_k$ vor und es ist

$$|t_{\alpha_i^x}^x - t_{\alpha_i^x}| \leqq |t_{\alpha_i^x}^x - t_{\alpha_i^x}^0| + |t_{\alpha_i^x}^0 - t_{\alpha_i^x}| < 2 \cdot \frac{1}{2m_i^x} = \frac{1}{m_i^x}.$$

Daraus folgt, daß die Koordinaten von ξ allen die Umgebung $U^{m_1^x, m_2^x, \ldots, m_k^x}_{a_1^x, a_2^x, \ldots, a_k^x}(x)$ bestimmenden Bedingungen genügen und also

$$\xi \subset U^{m_1^x, m_2^x, \ldots, m_k^x}_{a_1^x, a_2^x, \ldots, a_k^x} \subset G$$

ist.

Wir wollen jetzt zeigen, daß die eineindeutige Abbildung

$$R^* = f(R)$$

in beiden Richtungen stetig ist.

Der Beweis vollzieht sich in zwei Schritten.

I. Es sei $U(x_0^*) = V^{n_1 n_2 \ldots n_k}_{a_1 a_2 \ldots a_k}(x_0^*)$ beliebig gegeben. Man betrachte ein $X_{a_i}(x_0)$. Wenn $\varphi_{a_i}(x_0) = \psi_{a_i}(x_0)$ ist, so gibt es (wie bei der Definition der Funktionen φ und ψ erwähnt wurde) eine $U(x_0)$ in R — nennen wir sie W_{a_i} — derart, daß für sämtliche $x \subset W_{a_i}(x_0)$

$$\varphi_{a_i}(x) = \psi_{a_i}(x) = \varphi_{a_i}(x_j) = \psi_{a_i}(x_0)$$

ist, es ist m. a. W.

$$X_{a_i}(x) = X_{a_i}(x_0)$$

und also um so mehr

$$X_{a_i}(x) \subset S\left(X_{a_i}(x_0); \frac{1}{n_i}\right)$$

für alle $x \subset W_{a_i}(x_0)$.

Wenn aber $X_{a_i}(x_0)$ die ganze Einheitsstrecke ist, so ist für jeden Punkt $x \subset R$

$$X_{a_i}(x) \subset S\left(X_{a_i}(x_0); \frac{1}{n_i}\right)$$

und wir setzen dementsprechend $W_{a_i}(x_0) = R$.

Es sei jetzt $U(x_0)$ eine im Gebiete

$$W_{a_1}(x_0) \cdot W_{a_2}(x_0) \cdot \ldots \cdot W_{a_k}(x_0)$$

enthaltene Umgebung; offenbar ist dann für jeden Punkt $x \subset U(x_0)$ und jedes $i = 1, 2, \ldots, k$

$$X_{a_i}(x) \subset S\left(X_{a_i}(x_0); \frac{1}{n_i}\right);$$

diese Tatsache bedeutet aber gerade, daß

$$f(x) \subset V^{n_1 n_2 \ldots n_k}_{a_1 a_2 \ldots a_k}(x_0^*),$$

sobald $x \subset U(x_0)$.

II. Es sei eine beliebige $U(x_0) = U_a$ gegeben; um die Stetigkeit der Umkehrungsfunktion $f^{-1}(x^*)$ zu beweisen, genügt es zu zeigen, daß für jeden Punkt x^* von R^*, der in $V_a^2(x_0^*)$ liegt, $f^{-1}(x^*) \subset U_a$ ist. Nun ist aber $\varphi_a(x_0) = \psi_a(x_0) = 0$, also besteht $X_a(x_0)$ aus dem einzigen Punkte 0. Wenn $f^{-1}(x^*) = x \subset R - U_a$ wäre, wäre $\psi_a(x) = 1$, also könnte auch x^* definitionsgemäß nicht zu $V_a^2(x_0^*)$ gehören.

Die Homöomorphie zwischen R und R^* ist hiermit bewiesen. Da $R^* \subset R_\tau^R$ ist, haben wir noch zu zeigen, daß R_τ^R ein absolut abgeschlossener Raum ist. Dazu brauchen wir aber folgende zwei Bemerkungen zu machen:

1. Es gibt für jeden Punkt $x_0 = \{t_1^0, t_2^0, \ldots, t_\alpha^0, \ldots\}$ von R_τ abzählbare Punktmengen A, die zu diesem Punkte konvergieren, und zwar solche, daß unter den Koordinaten der Punkte von A weder 0 noch 1 vorkommt. Wir brauchen in der Tat z. B. die Punktmenge $x_n = \{t_1^{(n)}, t_2^{(n)}, \ldots, t_\alpha^{(n)}, \ldots\}$ $(n = 1, 2, \ldots)$ zu nehmen mit $t_\alpha^n = t_\alpha^0 - \frac{t_\alpha^0}{n+1}$, falls $t_\alpha^0 \neq 0$ und $t_\alpha^{(n)} = \frac{1}{n+1}$, falls $t_\alpha^{(0)} = 0$. Von einer solchen Folge sagen wir, sie konvergiere regelmäßig gegen x_0.

Sind alle Koordinaten eines Punktes x_0 von 0 und 1 verschieden, so ist $x_0 \subset R - \Xi$ und somit ist $\varphi(x_0) = x_0$; wenn x_n regelmäßig gegen x_0 konvergiert, so konvergiert auch $\varphi(x_n)$ gegen $\varphi(x_0)$ (weil ja auch $\varphi_n(x_n) = x_n$ ist.

2. Die den Raum R_τ^R bestimmende Zerlegung von R_τ ist im allgemeinen nicht stetig [14]). Es kann also vorkommen, daß für eine Umgebung $U(x_0^*)$ in R_τ^R die Menge $\varphi^{-1}(U)$ die Menge $X_0 = \varphi^{-1}(x_0^*)$ nicht im Innern enthält; dagegen existiert ein Teilgebiet G von R_τ von der Eigenschaft, daß

$$\varphi^{-1}(\overline{U(x_0^*)}) \supset G \supset X_0$$

ausfällt. Es genügt in der Tat, für G das der Definition von $U(x_0^*)$ zugrunde gelegtes Gebiet [15]) zu wählen. Es sei x ein beliebiger Punkt von G, $x_1, x_2, \ldots, x_n, \ldots$ eine gegen x regelmäßig konvergierende Folge. Da $\varphi(x_n) = x_n$ ist, so ist $\varphi(x_n) \subset U(x_0^*)$, also $\varphi(x) \subset \overline{U(x_0^*)}$.

Um jetzt die absolute Abgeschlossenheit des Raumes R_τ^R zu beweisen, genügt es nach einem Satze von Alexandroff und Urysohn [16]) zu zeigen, daß man aus jedem den Raum R_τ^R überdeckenden System von Umgebungen $\{U\}$ ein endliches Teilsystem $U_1, U_2, \ldots, U_N$ so wählen kann, daß

$$\overline{U}_1 + \overline{U}_2 + \ldots + \overline{U}_n = R_\tau^R$$

ist. Man betrachte zu diesem Zwecke die der Definition von $\{U\}$ zugrunde gelegten [17]) Gebiete G in R_τ. Da R_τ bikompakt ist und die $\{G\}$ eine

---

[14]) Aus dem Satze II und dem unter [18]) zitierten Satz von Alexandroff folgt vielmehr, daß unsere Zerlegung nur dann stetig sein kann, wenn der gegebene Raum regulär ist.

[15]) Alexandroff, a. a. O. § 3, S. 537.

[16]) Alexandroff und Urysohn, Zur Theorie der topologischen Räume, Math. Annalen 92 (1924), S. 262, Satz II.

[17]) Alexandroff a. a. O. § 3, S. 537.

Überdeckung von R_τ bilden, lassen sich endlichviele unter den $\{G\}$ — etwa $G_1, G_2, \ldots, G_N$ — so wählen, daß

$$G_1 + G_2 + \ldots + G_N = R_\tau$$

ist. Dann ist

$$\varphi(G_1) + \varphi(G_2) + \ldots + \varphi(G_N) = R_\tau^R;$$

nach dem soeben Bewiesenen ist aber $\varphi(G) \subset \overline{U}$, so daß

$$\overline{U}_1 + \overline{U}_2 + \ldots + \overline{U}_N$$

erst recht den ganzen Raum R_τ^R ausfüllt.

Dadurch ist die absolute Abgeschlossenheit von R_τ^R und mithin auch der Satz III vollständig bewiesen.

Ich möchte hervorheben, daß die von Alexandroff und Urysohn seit langem gestellte Frage, *ob jeder topologische Raum als überalldichte Teilmenge eines absolut abgeschlossenen Raumes betrachtet werden kann*, durch den soeben bewiesenen Satz keine Antwort erhalten hat[18]), da man nicht weiß, ob die Menge $\overline{R}^* \subset R_\tau^R$, als Raum betrachtet, absolut abgeschlossen ist. Auch die Frage, ob man alle topologischen Räume, die Basen von der Mächtigkeit τ besitzen, in *einen und denselben absolut abgeschlossenen Raum einbetten kann*, bleibt unentschieden.

---

[18]) Eine abgeschlossene Teilmenge eines absolut abgeschlossenen Raumes braucht nicht (wie elementare Beispiele zeigen) absolut abgeschlossen zu sein (siehe Alexandroff und Urysohn, Mémoire sur les espaces topologiques compacts, Verh. Kon. Ak. Amsterdam, Deel XIV, No. 1 (1929)).

(Eingegangen am 6. 1. 1929.)

Zur Topologie der Abbildungen von Mannigfaltigkeiten.

Zweiter Teil[1]).

Klasseninvarianten von Abbildungen.

Von

Heinz Hopf in Berlin.

Brouwer hat in seiner grundlegenden Arbeit „Über Abbildung von Mannigfaltigkeiten“ [2]) den Abbildungsgrad durch den Beweis des folgenden Satzes eingeführt: „*Wenn eine zweiseitige, geschlossene, gemessene n-dimensionale Mannigfaltigkeit* μ *auf eine gemessene n-dimensionale Mannigfaltigkeit* μ' *eindeutig und stetig abgebildet wird, so existiert eine bei stetiger Modifizierung der Abbildung sich nicht ändernde endliche ganze Zahl* c *mit der Eigenschaft, daß die Bildmenge von* μ *jedes Teilgebiet von* μ' *im ganzen* c *Male positiv überdeckt. Ist* μ' *einseitig oder offen, so ist* c *stets gleich Null.*“ Dabei ist, wie aus dem Beweis des Satzes hervorgeht, der „*Abbildungsgrad*“ c im „algebraischen“ Sinne als Anzahl der positiven Bedeckungen aufzufassen, d. h. er ist die Anzahl der positiven Bedeckungen vermindert um die Anzahl der negativen Bedeckungen. Der Grad ändert sich, wie der Satz besagt, nicht, wenn die gegebene Abbildung f stetig modifiziert wird, er ist also eine Invariante der durch f bestimmten „Abbildungsklasse“; die ohne Berücksichtigung von Vorzeichen bestimmte Anzahl der glatten, d. h. eineindeutigen, Bedeckungen eines Gebietes durch die Bildmenge dagegen ist weder auf μ' noch in der Klasse konstant; über sie kann man von vornherein nur aussagen, daß sie niemals kleiner als $|c|$ sein kann. So entsteht folgende Frage: „*Gibt es in der durch* f *bestimmten Abbildungsklasse eine Ab-*

---

[1]) Erster Teil: Neue Darstellung der Theorie des Abbildungsgrades für topologische Mannigfaltigkeiten, Math. Annalen **100** (1928). Im folgenden als „Teil I“ zitiert.

[2]) Brouwer, Über Abbildung von Mannigfaltigkeiten, Math. Annalen **71** (1911).

bildung f', bei der die Bildmenge ein Gebiet von μ' nicht nur „algebraisch“, sondern auch „geometrisch“, d. h. in bezug auf eine Abzählung, die auf Vorzeichen keine Rücksicht nimmt, $|c|$-mal bedeckt? Die Beantwortung dieser Frage, auf deren Bedeutung ich durch Herrn Alexandroff hingewiesen worden bin, bildet das eigentliche Ziel dieser Arbeit.

Die Frage ist zu bejahen, wenn μ und μ' orientierbar sind. Dies wird unten für den Fall $n \neq 2$ bewiesen, während der Fall $n = 2$, der sich unserer Beweismethode entzieht, in demselben Sinne von H. Kneser erledigt worden ist[3]). Wenn also beide Mannigfaltigkeiten orientierbar sind, so läßt sich der Grad seinem Betrage nach charakterisieren als *die innerhalb der vorgelegten Abbildungsklasse mögliche Mindestzahl glatter Bedeckungen eines Gebietes von μ' durch die Bildmenge,* — eine Tatsache, die die von Brouwer in einer späteren Arbeit[4]) bewiesene topologische Invarianz des Grades in Evidenz setzt.

Diese Übereinstimmung des Grades mit der soeben genannten Mindestzahl besteht aber offenbar nicht, wenn μ' nicht orientierbar ist; denn dann ist ja nach dem am Anfang zitierten Brouwerschen Satz immer $c = 0$, während die fragliche Mindestzahl z. B. bei der Abbildung einer Kugel auf eine projektive Ebene, bei der das Bild der ersteren als unverzweigte zweiblättrige Überlagerungsfläche über letzterer liegt, wie man leicht erkennt, 2 ist. Hier ist also die Aufgabe zu lösen, diese Mindestzahl, die ihrer Definition nach eine Klasseninvariante ist, mittels Eigenschaften der vorgelegten Abbildung f zu bestimmen, und dieselbe Aufgabe tritt auf, wenn auch μ nicht orientierbar, der Grad also gar nicht definiert ist. Man wird versuchen, für diese Fälle eine dem Grade ähnliche Zahl zu erklären, von der sich nachträglich zeigen läßt, daß sie für die in der Klasse möglichen Bedeckungszahlen die Minimaleigenschaft hat, die nach dem oben Gesagten im Fall orientierbarer Mannigfaltigkeiten dem Grade zukommt. Dies wird unten durchgeführt; die Klasseninvariante, die die Rolle des Grades übernimmt, wird als „Absolutgrad“ bezeichnet und ist im Falle geschlossener orientierbarer Mannigfaltigkeiten mit dem absoluten Betrag des Grades identisch; für ihre Definition ist unter anderem eine feinere Einteilung der Abbildungen nicht orientierbarer Mannigfaltigkeiten in „orientierbare“ und „nicht orientierbare“ Abbildungen notwendig, die das Verhalten der Bilder derjenigen geschlossenen Wege berücksichtigt, bei deren Durchlaufung sich die Orientierung umkehrt, und die in ähnlicher Weise gelegentlich ebenfalls schon von Brouwer vorgenommen worden ist[5]). Der Beweis der Übereinstimmung des

---

[3]) H. Kneser, Glättung von Flächenabbildungen, Math. Annalen **100** (1928).

[4]) Brouwer, Über Jordansche Mannigfaltigkeiten, § 6, Math. Annalen **71** (1911).

[5]) Brouwer, Aufzählung der Abbildungsklassen endlichfach zusammenhängender Flächen, Math. Annalen **82** (1921); besonders S. 284.

Absolutgrades mit der oben besprochenen minimalen Bedeckungszahl versagt auch hier für den Fall $n = 2$; aber auch diese Lücke ist von H. Kneser ausgefüllt worden[6]).

Nun ist aber der Bereich der Gebilde, für deren Abbildungen der Grad definiert ist, durch die *geschlossenen*, orientierbaren oder nicht orientierbaren, Mannigfaltigkeiten keineswegs erschöpft, und gerade bei manchen Anwendungen spielen die Abbildungen von *berandeten* oder von den aus diesen durch Weglassung des Randes entstehenden *offenen* Mannigfaltigkeiten eine wesentliche Rolle; es braucht wohl nur daran erinnert zu werden, daß Brouwers Beweis der Invarianz der Dimensionenzahl[7]) auf der Betrachtung des — dort allerdings noch nicht mit diesem Namen versehenen — „Grades" der Abbildung eines n-dimensionalen Würfels beruht. Die Theorie des Grades für die Abbildungen offener Mannigfaltigkeiten, auf die sich, wie schon angedeutet, die Betrachtung berandeter Mannigfaltigkeiten stets zurückführen läßt, ist im ersten Teil der vorliegenden Arbeit ausführlich dargestellt worden; der Grad ist hier nicht mehr auf μ', sondern nur in gewissen Gebieten konstant; ferner hat man die Gesamtheit der zulässigen „stetigen Modifizierungen" der Abbildungen einzuschränken, und zwar ungefähr in dem Sinne, daß man, wenn eine berandete Mannigfaltigkeit abgebildet wird, die Abbildung am Rande nicht ändern darf. Für diesen weiteren Bereich von Abbildungen ergibt sich nun ebenso wie bei geschlossenen Mannigfaltigkeiten die Frage nach der durch zulässige stetige Abänderungen erreichbaren Mindestzahl glatter Bedeckungen eines Gebietes, und der oben genannte Weg zur Antwort auf diese Frage ist auch hier gangbar: die Definition des „Absolutgrades" gilt von vornherein für die Abbildungen beliebiger, geschlossener oder offener, Mannigfaltigkeiten, und auch der Nachweis seines Übereinstimmens mit der in Frage stehenden Mindestzahl hat für $n \neq 2$ diesen allgemeinen Gültigkeitsbereich. Jedoch zeigt es sich jetzt, daß die Dimensionszahl $n = 2$ nicht nur unserer Beweismethode unzugänglich ist, sondern daß sie wirklich in bezug auf unsere Fragestellung eine Ausnahmerolle spielt: *es gibt Abbildungen offener Flächen, für welche die durch zulässige Abänderungen erreichbare Mindestzahl der glatten Bedeckungen eines Gebietes größer ist als der Absolutgrad* — im Gegensatz zu allen anderen Dimensionszahlen, wo die beiden Zahlen stets einander gleich sind. Für die Abbildungen offener Flächen bleibt unser Problem also ungelöst, und seine nähere Untersuchung scheint auf gruppentheoretische Fragen zu führen[8]).

---

[6]) H. Kneser, Die kleinste Bedeckungszahl innerhalb einer Klasse von Flächenabbildungen; erscheint in den Math. Annalen.

[7]) Brouwer, Beweis der Invarianz der Dimensionenzahl, Math. Annalen 70 (1911).

[8]) Vgl. Fußnote [30]).

Die bisher besprochene Mindestzahl läßt sich auch folgendermaßen schildern: Man faßt einen festen Punkt ξ von μ' und solche Abbildungen f aus der vorgelegten Klasse ins Auge, bei denen die Gleichung $f(x) = \xi$ nur „einfache" Lösungen x besitzt; dabei heißt ein Punkt x auf μ eine einfache Lösung, wenn es eine Umgebung von x gibt, in der die Gleichung $f(y) = \eta$ für jeden hinreichend nahe an ξ gelegenen Punkt η genau eine Lösung y hat; geometrisch gesprochen: es werden nur glatte Bedeckungen der Umgebung von ξ zugelassen. Gefragt wird nach einer Abbildung f, bei der die Anzahl dieser einfachen Lösungen möglichst klein ist. Bei dieser Formulierung drängt sich von selbst die Frage nach einer Abbildung der vorgelegten Klasse auf, für die die Anzahl der Lösungen von $f(x) = \xi$ schlechthin, ohne Rücksicht auf Einfachheit, möglichst klein ist, also nach einer Abbildung, die *die innerhalb der Klasse erreichbare Mindestzahl der Originalpunkte des Punktes* ξ liefert. Diese Mindestzahl ist ebenso wie die oben besprochene nach Definition eine Invariante der Abbildungsklasse und des Punktes ξ; für den Fall geschlossener Mannigfaltigkeiten μ ist sie wieder, wie man leicht sieht, von ξ unabhängig. Ihre Untersuchung ist in dieser Arbeit ein wichtiger Schritt auf dem Wege zu dem Absolutgrad und der mit diesem verknüpften Mindestzahl der glatten Bedeckungen.

Die somit gestellte Aufgabe, die in der Klasse mögliche Mindestzahl der Originale von ξ zu bestimmen, ist ihrem Wesen nach nahe verwandt mit der von J. Nielsen in Angriff genommenen Frage nach der innerhalb einer Klasse von Abbildungen von μ auf sich erreichbaren Mindestzahl von Fixpunkten*), also von Lösungen der Gleichung $f(x) = x$; und der Begriff, der bei uns im wesentlichen zum Ziele führt, ist einem Nielsenschen Grundbegriff nachgebildet: analog der durch Nielsen vorgenommenen Einteilung der Fixpunktmenge in „Fixpunktklassen" zerlegen wir die Originalmenge von ξ in „Schichten" und unterscheiden „wesentliche" und „unwesentliche" Schichten. Die Anzahl der wesentlichen Schichten ist eine Invariante der Klasse (und hängt, falls μ offen ist, von dem Punkte ξ ab); sie ist eine untere Schranke für die jetzt zu untersuchende Mindestzahl; es bleibt die Frage zu beantworten, ob sie gleich dieser Mindestzahl ist. Die Antwort fällt analog der Antwort auf die früher gestellte Frage nach der ersten von uns betrachteten Mindestzahl aus: *sie lautet „ja", falls* $n \neq 2$ *ist; dagegen gibt es Abbildungsklassen von Flächen — und hier sogar von geschlossenen Flächen —, für die die Mindestzahl der Originale eines Punktes größer als die wesentliche Schichten-*

*) J. Nielsen, Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen I, II. Acta mathematica **50**, **53** (1927, 1929). Man vergleiche besonders die Charakterisierung der Fixpunktklassen in I auf S. 289 unten.

zahl ist, und die Frage, wie die Mindestzahl allgemein zu bestimmen ist, bleibt offen.

Die Bestimmung der Anzahl der wesentlichen Schichten führt, wie es bei der Analogie mit den Nielsenschen Begriffen zu erwarten ist, auf die Betrachtung der Beziehung zwischen den Fundamentalgruppen der beiden Mannigfaltigkeiten, die durch die Abbildung vermittelt wird. *Ist μ geschlossen und der Grad (bzw. Absolutgrad) von 0 verschieden, so ist die Anzahl durch den genannten Gruppenhomomorphismus vollständig bestimmt: das Bild der Fundamentalgruppe von μ ist eine Untergruppe der Fundamentalgruppe von μ' mit endlichem Index; dieser Index ist die Zahl der wesentlichen Schichten und ein Teiler des Grades (bzw. Absolutgrades).*

Dies sind, in kurzen Zügen, die Probleme und Ergebnisse dieser Arbeit. Man sieht, daß die am wenigsten geklärten Punkte in unserem Problemkreis diejenigen sind, die sich auf Flächenabbildungen beziehen. Hier wird in zwei Richtungen weiterzuarbeiten sein: einerseits muß man die Abbildungsklassen der Flächen selbst, insbesondere der geschlossenen Flächen, noch genauer untersuchen, als es bisher geschehen ist, um z. B. die Mindestzahl der Originale eines Bildpunktes mittels Eigenschaften des durch die Abbildung bewirkten Gruppenhomomorphismus zu bestimmen (ein erstrebenswertes Ziel ist hier natürlich die Aufzählung aller Abbildungsklassen durch Angabe aller möglichen Homomorphismen von Flächengruppen[10]); andererseits wird man zu versuchen haben, den Fall $n = 2$ durch Modifizierung oder Verallgemeinerung der Fragestellung von seiner Ausnahmestellung zu befreien und in eine allgemeinere Theorie einzuordnen. In beiden Richtungen werden in zwei der Arbeit angefügten Anhängen kleine Vorstöße unternommen.

Inhaltsverzeichnis.

[10]) Daß die Aufzählung der Gruppenhomomorphismen im allgemeinen (nämlich dann, wenn μ' nicht die Kugel oder die projektive Ebene ist) mit der Aufzählung der Abbildungsklassen identisch ist, geht aus der unter [8]) genannten Arbeit von Brouwer (S. 286) hervor. Die unter [6]) genannte Arbeit von Kneser enthält wichtige Beiträge zur Durchführung dieser Aufzählung.

§ 1.

Vorbemerkungen über Fundamentalgruppe und Überlagerungsmannigfaltigkeiten [11]).

Unter einem „Weg" in der Mannigfaltigkeit M verstehen wir das eindeutige und stetige Bild einer Strecke, unter seinem Anfangs- und Endpunkt die Bilder des Anfangs- und Endpunkts der Strecke. Es ist klar, was unter dem zu einem Weg w inversen Weg w^{-1} sowie unter der Zusammensetzung $w_1 w_2$ zweier Wege zu verstehen ist. Ein Weg w heißt geschlossen, wenn Anfangs- und Endpunkt zusammenfallen; ein geschlossener Weg läßt sich also als eindeutiges und stetiges Bild einer Kreisperipherie auffassen. Er ist dann und nur dann (auf einen Punkt) „zusammenziehbar", wenn sich diese Abbildung zu einer Abbildung der ganzen Kreisscheibe ergänzen läßt. Zwei geschlossene Wege w_1, w_2 mit gemeinsamem Anfangs- und Endpunkt heißen „äquivalent", wenn der Weg $w_1 w_2^{-1}$ zusammenziehbar ist; wir schreiben dann $w_1 \equiv w_2$. Aus $w_1 \equiv w_2$, $w_1 \equiv w_3$ folgt leicht $w_2 \equiv w_3$; die geschlossenen Wege mit gemeinsamem Anfangs- und Endpunkt x lassen sich also in Klassen zusammenfassen. Zwei geschlossene Wege durch x sind dann und nur dann unter Festhaltung ihres gemeinsamen Anfangs- und Endpunkts ineinander deformierbar, wenn sie zu einer Klasse gehören. Die Klassen bilden bezüglich der erwähnten Zusammensetzung eine Gruppe $\mathfrak{F}_x$. Ist y ein von x verschiedener Punkt, so ist die entsprechend definierte Gruppe $\mathfrak{F}_y$ mit $\mathfrak{F}_x$ isomorph; der Isomorphismus ist, wenn g_x ein geschlossener Weg durch x, v ein fester Weg von x nach y ist, dadurch gegeben, daß man dem Element g_x von $\mathfrak{F}_x$ das Element

---

[11]) Dieser Paragraph, insbesondere sein erster Absatz, ist nur eine Zusammenstellung von Tatsachen, die ziemlich allgemein bekannt sein dürften. — Literatur: Poincaré, Analysis Situs, §§ 12, 13, Journ. École Polytechn. (2) 1 (1895). — Kerékjártó, Vorlesungen über Topologie (Berlin 1923), V. Abschn., § 2. — J. Nielsen, wie unter [9]). — Reidemeister, Fundamentalgruppe und Überlagerungsräume, Nachr. Ges. d. Wiss. Göttingen, Math.-phys. Klasse, 1928. — Schreier, Die Verwandtschaft stetiger Gruppen im großen, § 1, Abhandl. Math. Sem. d. Hamburgischen Universität 5 (1927). — Weyl, Die Idee der Riemannschen Fläche (Leipzig-Berlin 1913), § 9.

$g_y = v^{-1} g_x v$ von $\mathfrak{F}_y$ zuordnet. Die Willkür von v hat zur Folge, daß dieser Isomorphismus nur bis auf innere Automorphismen bestimmt ist: wenn man nämlich v durch $\bar{v}$ ersetzt und $\bar{v}^{-1} v = h$ setzt, so wird dem Element g_x nicht $g_y = v^{-1} g_x v$, sondern $\bar{g}_y = \bar{v}^{-1} g_x \bar{v} = h v^{-1} g_x v h^{-1} = h g_y h^{-1}$ zugeordnet. Die abstrakte Gruppe $\mathfrak{F}$, von der $\mathfrak{F}_x, \mathfrak{F}_y, \ldots$ Realisationen sind, heißt die „Fundamentalgruppe" von M.

Zu jeder Untergruppe $\mathfrak{U}$ von $\mathfrak{F}$ wird folgendermaßen die „zu $\mathfrak{U}$ gehörige Überlagerungsmannigfaltigkeit von M" konstruiert:

x sei ein fester Punkt in M, $\mathfrak{U}_x$ eine der Untergruppen von $\mathfrak{F}_x$, die $\mathfrak{U}$ entsprechen (die also alle die Form $h\,\mathfrak{U}_x h^{-1}$ haben). Wir betrachten zunächst die Wege mit dem Anfangspunkt x und einem festen Endpunkt y und nennen zwei solche Wege w_1, w_2 „äquivalent $\operatorname{mod} \mathfrak{U}_x$", geschrieben: $w_1 \equiv w_2 \pmod{\mathfrak{U}_x}$, wenn $w_1 w_2^{-1} \subset \mathfrak{U}_x$ ist. Ist

$$w_1 \equiv w_2 \pmod{\mathfrak{U}_x}, \qquad w_1 \equiv w_3 \pmod{\mathfrak{U}_x},$$

so ist also

$$w_1 w_2^{-1} \subset \mathfrak{U}_x, \qquad w_2 w_1^{-1} \subset \mathfrak{U}_x, \qquad w_1 w_3^{-1} \subset \mathfrak{U}_x,$$

mithin

$$w_2 w_1^{-1} w_1 w_3^{-1} \equiv w_2 w_3^{-1} \subset \mathfrak{U}_x, \text{ [12]}$$

d. h.

$$w_2 \equiv w_3 \pmod{\mathfrak{U}_x}.$$

Die von x nach y laufenden Wege lassen sich also in „Äquivalenzklassen $\operatorname{mod} \mathfrak{U}_x$" einteilen.

Ist w_0 ein fester Weg von x nach y und sind w_1, w_2 irgend zwei Wege von x nach y, so ist, wenn wir die geschlossenen Wege $w_1 w_0^{-1}$, $w_2 w_0^{-1}$ mit g_1, g_2 bezeichnen, $w_1 w_2^{-1} \equiv g_1 g_2^{-1}$; w_1, w_2 gehören also dann und nur dann zur selben Klasse $\operatorname{mod} \mathfrak{U}_x$, wenn $g_1 g_2^{-1} = u \subset \mathfrak{U}_x$, $g_1 \equiv u g_2$ ist, d. h. wenn g_1 und g_2 derselben Restklasse („Nebengruppe") $\operatorname{mod} \mathfrak{U}_x$ angehören. Den Äquivalenzklassen der Wege von x nach y sind also — durch Vermittlung des willkürlich ausgezeichneten Weges w_0 — eineindeutig Restklassen von $\mathfrak{F}_x \operatorname{mod} \mathfrak{U}_x$ zugeordnet. Dabei kommt jede dieser Restklassen vor; denn ist g irgendein Weg aus $\mathfrak{F}_x$, so entspricht bei der Zuordnung der Äquivalenzklasse von $w = g w_0$ die Restklasse, der g angehört. Somit sehen wir: Dem System der Äquivalenzklassen der Wege von x nach y entspricht — nach Auszeichnung von w_0 — eineindeutig das System der Restklassen, in das $\mathfrak{F}_x \operatorname{mod} \mathfrak{U}_x$, oder, was dasselbe ist, in das $\mathfrak{F} \operatorname{mod} \mathfrak{U}$ zerfällt; insbesondere ist die — endliche oder unendliche —

---

[12]) Man ziehe $w_1^{-1} w_1$ unter Festhaltung des Endpunktes von w_1 auf diesen zusammen.

Anzahl der Äquivalenzklassen gleich der Anzahl j der Restklassen, die, wie üblich, der „Index" von $\mathfrak{U}$ in $\mathfrak{F}$ heißen möge[13]).

Wir lassen nun, immer unter Festhaltung von x, den Punkt y die Mannigfaltigkeit M durchlaufen und bekommen so eine Menge $\mathfrak{M}$ von Äquivalenzklassen, die so auf M abgebildet ist, daß jeder Punkt y Bild von j Klassen ist. Wir machen $\mathfrak{M}$ durch geeignete Definition eines Umgebungsbegriffs zu einer Mannigfaltigkeit:

Um jeden Punkt y zeichnen wir eine feste euklidische Umgebung U_y aus. Sind $w_1(y) = w_1$, $w_2(y) = w_2$ zwei zu einer Klasse W gehörige Wege von x nach y und sind v_1, v_2 zwei innerhalb U_y von y nach einem Punkt y' von U_y verlaufende Wege, so ist $v_1 v_2^{-1}$ zusammenziehbar, also ist auch $w_1 v_1 v_2^{-1} w_1^{-1}$ zusammenziehbar; es ist also $w_1 v_1 v_2^{-1} w_1^{-1} \subset \mathfrak{U}_x$, $w_1 w_2^{-1} \subset \mathfrak{U}_x$, mithin, wenn wir $w_1(y)v_1 = w_1'(y') = w_1'$, $w_2(y)v_2 = w_2'(y') = w_2'$ setzen,

$$w_1' w_2'^{-1} = w_1 v_1 v_2^{-1} w_1^{-1} w_1 w_2^{-1} \subset \mathfrak{U}_x, \quad \text{d. h.} \quad w_1' \equiv w_2' \pmod{\mathfrak{U}_x}.$$

Die somit durch die Klasse W der $w_i(y)$ bestimmte Klasse W' der $w_i'(y')$ nennen wir die zu W „benachbarte" Klasse der Wege von x nach y'. Ersetzt man die Wege $w_i(y)$ durch einen zu einer anderen Klasse $\overline{W}$ gehörigen Weg $\bar{w}(y)$, so geht die zu $\overline{W}$ benachbarte Klasse $\overline{W}'$ der nach y' führenden Wege aus der Klasse W' durch linksseitige Multiplikation mit $\bar{w}w^{-1}$ hervor, wobei $w \subset W$ ist; da $\bar{w}w^{-1}$ nicht zu $\mathfrak{U}_x$ gehört, ist $\overline{W}'$ von $\overline{W}$ verschieden. Zu voneinander verschiedenen nach y führenden Wegeklassen sind also voneinander verschiedene noch y' führende Wegeklassen benachbart.

Wir setzen nun fest, daß eine Menge von Klassen $W'(y')$ eine Umgebung der Klasse $W(y)$ heißt, wenn erstens die Punkte y' eine in U_y enthaltene euklidische Umgebung von y bilden, und wenn zweitens die $W'(y')$ zu $W(y)$ benachbart sind. Man überzeugt sich leicht von folgenden Tatsachen: Die Umgebungsdefinition erfüllt die Hausdorffschen Axiome, $\mathfrak{M}$ ist also zu einem topologischen Raum gemacht worden; die (durch die Endpunkte der Wege bestimmte) Abbildung von $\mathfrak{M}$ auf M ist stetig; zu jedem der in M ausgezeichneten Gebiete U_y gibt es in $\mathfrak{M}$ genau j Gebiete $U_y^1, U_y^2, \ldots$, die auf U_y abgebildet sind; diese Beziehung zwischen U_y^i und U_y ist eineindeutig und beiderseits stetig. Da somit die in $\mathfrak{M}$ definierten Umgebungen topologische Bilder euklidischer Gebiete sind, ist $\mathfrak{M}$ eine Mannigfaltigkeit.

---

[13]) Der Index ist unabhängig davon, ob man rechtsseitige oder linksseitige Restklassen betrachtet; denn bei dem Automorphismus von $\mathfrak{F}$, der entsteht, wenn man jedes Element durch sein Inverses ersetzt, werden die beiden Systeme der rechts- und linksseitigen Restklassen eineindeutig aufeinander abgebildet.

Wir waren ausgegangen von einer Untergruppe $\mathfrak{U}$ von $\mathfrak{F}$ und hatten der weiteren Betrachtung eine der Untergruppen von $\mathfrak{F}_x$ zugrunde gelegt, die bei dem zwischen $\mathfrak{F}$ und $\mathfrak{F}_x$ bestehenden Isomorphismus $\mathfrak{U}$ entsprechen; wir hätten statt dieser Untergruppe $\mathfrak{U}_x$ auch eine zu $\mathfrak{U}_x$ ähnliche Untergruppe $\overline{\mathfrak{U}}_x = h\,\mathfrak{U}_x h^{-1}$ zugrunde legen können. Dann wären wir statt zu $\mathfrak{M}$ zu einer anderen Mannigfaltigkeit $\overline{\mathfrak{M}}$ gelangt. Jedoch läßt sich $\overline{\mathfrak{M}}$ mit $\mathfrak{M}$ ohne Abänderung der zwischen $\overline{\mathfrak{M}}$ und M bzw. zwischen $\mathfrak{M}$ und M bestehenden Beziehung dadurch identifizieren, daß man den Punkt von $\mathfrak{M}$, zu dem die Wege $w_1, w_2, \ldots$ gehören, als identisch mit dem Punkt von $\overline{\mathfrak{M}}$ betrachtet, zu dem die Wege $h w_1, h w_2, \ldots$ gehören. Dies ist möglich, denn $w_1 \equiv w_2 \pmod{\mathfrak{U}_x}$ ist gleichbedeutend mit $h w_1 \equiv h w_2 \pmod{\overline{\mathfrak{U}}_x}$. Faßt man also $\mathfrak{M}$, $\overline{\mathfrak{M}}$ als Realisationen einer abstrakten Mannigfaltigkeit auf, so ist diese sowie ihre Abbildung auf M unabhängig davon, welche der $\mathfrak{U}$ entsprechenden Untergruppen von $\mathfrak{F}_x$ wir zugrunde gelegt haben.

Hätten wir schließlich statt des Punktes x einen Punkt z von M zugrunde gelegt, so wären wir statt zu $\mathfrak{M} = \mathfrak{M}_x$ zu einer Mannigfaltigkeit $\mathfrak{M}_z$ gelangt. Aber $\mathfrak{M}_x$ und $\mathfrak{M}_z$ lassen sich ohne Störung ihrer Beziehungen zu M dadurch identifizieren, daß man den Punkt von $\mathfrak{M}_x$, zu dem die Wege $w_1, w_2, \ldots$ gehören, als identisch betrachtet mit dem Punkt von $\mathfrak{M}_z$, zu dem die Wege $v w_1, v w_2, \ldots$ gehören, wobei v ein fester Weg von z nach x ist; dies ist möglich, denn $w_1 \equiv w_2 \pmod{\mathfrak{U}_x}$ ist gleichbedeutend mit $v w_1 \equiv v w_2 \pmod{v\,\mathfrak{U}_x v^{-1}}$, und $\mathfrak{U}_z = v\,\mathfrak{U}_x v^{-1}$ ist eine Untergruppe von $\mathfrak{F}_z$, die $\mathfrak{U}$ entspricht. $\mathfrak{M}_x$ und $\mathfrak{M}_z$ können also wieder als Realisationen einer abstrakten Mannigfaltigkeit aufgefaßt werden, die samt ihrer Abbildung auf M von der Wahl der Punkte x oder z nicht abhängt.

Das Ergebnis ist: *Zu jeder Untergruppe $\mathfrak{U}$ der Fundamentalgruppe $\mathfrak{F}$ gehört eine wohlbestimmte abstrakte Mannigfaltigkeit $M_{\mathfrak{U}}$, die durch eine eindeutige und stetige Abbildung φ so auf M bezogen ist, daß auf jedes hinreichend kleine — nämlich in einem U_y enthaltene — Gebiet von M genau j Gebiete von $M_{\mathfrak{U}}$ topologisch abgebildet sind; dabei ist j der Index von $\mathfrak{U}$ in $\mathfrak{F}$. $M_{\mathfrak{U}}$ heißt die zu $\mathfrak{U}$ gehörige „Überlagerungsmannigfaltigkeit“ von M, j die Anzahl der „Schichten“ von $M_{\mathfrak{U}}$ über M.*

Wir betrachten nun noch die durch φ vermittelte Beziehung zwischen den Wegen auf $M_{\mathfrak{U}}$ und den Wegen auf M. Sei x ein Punkt in M, $\mathfrak{U}_x$ eine der $\mathfrak{U}$ entsprechenden Untergruppen von $\mathfrak{F}_x$, $\bar{x}$ der Punkt in $M_{\mathfrak{U}}$, der durch die von x nach x führenden Wege, welche zu $\mathfrak{U}_x$ gehören, also z. B. durch den nur aus dem Punkt x bestehenden Weg, definiert ist. Läuft in M ein Punkt x_t in stetiger Abhängigkeit von t in $x = x_0$ beginnend und ist w_t der bis zum Wert t durchlaufene Weg, $\bar{x}_t$ der durch w_t in $M_{\mathfrak{U}}$ bestimmte Punkt, so bilden diese Punkte, da φ in jedem hinreichend kleinen Gebiet von M eindeutig und stetig umkehrbar ist, einen

Weg $\bar{w}$ in $M_{\mathfrak{U}}$, und dessen Anfangspunkt ist $\bar{x}$, da dieser Punkt, wie eben bemerkt, infolge seiner Definition dem Weg w_0 entspricht. Hat man also in $M_{\mathfrak{U}}$ einen in $\bar{x}$ beginnenden Weg $\bar{w}$, so entspricht seinem Bild $w = \varphi(\bar{w})$ gerade der Endpunkt von $\bar{w}$.

Ist $\bar{w}$ geschlossen, so ist wegen der Eindeutigkeit von φ auch $w = \varphi(\bar{w})$ geschlossen; nach dem eben Gesagten entspricht dem Weg w der Endpunkt $\bar{x}$ von $\bar{w}$; folglich gehört w ebenso wie w_0 zu $\mathfrak{U}_x$, da, wie wir früher sahen, Wegen, die zu verschiedenen Restklassen mod $\mathfrak{U}_x$ gehören, verschiedene Punkte in $M_{\mathfrak{U}}$ entsprechen. Ist andererseits w ein zu $\mathfrak{U}_x$ gehöriger geschlossener Weg, so ist der seinen Teilbögen w_t entsprechende Weg in $M_{\mathfrak{U}}$ geschlossen, da $\bar{x}$ ja gerade durch die von x nach x führenden, zu $\mathfrak{U}_x$ gehörigen Wege definiert ist. Somit bildet φ die geschlossenen Wege durch $\bar{x}$ auf alle die und nur die geschlossenen Wege durch x ab, die zu $\mathfrak{U}_x$ gehören.

Ist der geschlossene Weg $\bar{w}$ in $M_{\mathfrak{U}}$ zusammenziehbar, so ist er das Randbild bei einer Abbildung $f(K)$ einer Kreisscheibe K; dann ist $w = \varphi(\bar{w})$ das Randbild bei der Abbildung $\varphi f(K)$ von K, also auch zusammenziehbar. Ferner erhält φ infolge ihrer Eindeutigkeit und Stetigkeit die Zusammensetzung geschlossener Wege, d. h. es ist $\varphi(\bar{w}_1 \bar{w}_2) = \varphi(\bar{w}_1)\,\varphi(\bar{w}_2)$; insbesondere ist also mit $\bar{w}_1 \bar{w}_2^{-1}$ auch $\varphi(\bar{w}_1)\,\varphi(\bar{w}_2)^{-1}$ zusammenziehbar, d. h. äquivalente Wege in $M_{\mathfrak{U}}$ werden auf äquivalente Wege in M abgebildet. Mithin bildet φ die Fundamentalgruppe $\bar{\mathfrak{U}}$ von $M_{\mathfrak{U}}$ homomorph auf $\mathfrak{U}_x$ ab.

Dieser Homomorphismus ist eineindeutig, also ein Isomorphismus. Um das zu beweisen, muß gezeigt werden, daß verschiedene Klassen von $\bar{\mathfrak{U}}$ auf verschiedene Klassen von $\mathfrak{U}_x$ abgebildet werden. Hierzu genügt der Nachweis der folgenden Behauptung: Ist $\varphi(\bar{w})$ zusammenziehbar, so ist auch $\bar{w}$ zusammenziehbar.

Sei also $w = \varphi(\bar{w})$ das Bild der Peripherie p bei der Abbildung $f(K)$ einer Kreisscheibe K und sei dabei $\mathfrak{x}$ der Punkt auf p, der durch f in den Anfangs- und Endpunkt x von w abgebildet wird. Ist $\mathfrak{w}_1$ ein Weg in K von $\mathfrak{x}$ nach einem Punkt $\mathfrak{y}$ in K, so entspricht dem in M verlaufenden Weg $f(\mathfrak{w}_1)$ ein bestimmter Punkt $\bar{y}$ in $M_{\mathfrak{U}}$; ist $\mathfrak{w}_2$ ein zweiter Weg in K von $\mathfrak{x}$ nach $\mathfrak{y}$, so ist $\mathfrak{w}_1 \mathfrak{w}_2^{-1}$ in K zusammenziehbar, also $f(\mathfrak{w}_1 \mathfrak{w}_2^{-1}) = f(\mathfrak{w}_1)\,f(\mathfrak{w}_2)^{-1}$ in M zusammenziehbar, also gewiß $f(\mathfrak{w}_1)\,f(\mathfrak{w}_2)^{-1} \subset \mathfrak{U}_x$; dann entspricht aber den Wegen $f(\mathfrak{w}_1)$ und $f(\mathfrak{w}_2)$ derselbe Punkt $\bar{y}$ in $M_{\mathfrak{U}}$. Die durch die Wege $\mathfrak{w}$ und $f(\mathfrak{w})$ vermittelte Abbildung F von K auf $M_{\mathfrak{U}}$ ist also eindeutig und infolge der Stetigkeit von f und der stetigen Umkehrbarkeit im Kleinen von φ stetig. Die Randabbildung $F(p)$ liefert $\bar{w}$; denn die auf p verlaufenden Bögen werden durch f auf die Teilbögen w_t von w abgebildet, und diesen entspricht, wie

wir oben sahen, der Weg $\bar{w}$. — Damit ist die Behauptung und somit der folgende Satz bewiesen:

$\mathfrak{U}$ *ist die Fundamentalgruppe von* $M_{\mathfrak{U}}$.

Betrachten wir neben $\mathfrak{U}$ eine zu $\mathfrak{U}$ ähnliche Untergruppe von $\mathfrak{F}$: $\mathfrak{U}' = h\mathfrak{U}h^{-1}$, wobei h ein Element von $\mathfrak{F}$ ist. Sind wieder w_1, w_2 von x nach y führende Wege, die mod $\mathfrak{U}$ äquivalent sind, so sind hw_1, hw_2 mod $\mathfrak{U}'$ äquivalent und umgekehrt, denn $w_1 w_2^{-1} \subset \mathfrak{U}$ ist gleichbedeutend mit $(hw_1)(hw_2)^{-1} = hw_1w_2^{-1}h^{-1} \subset \mathfrak{U}'$. Wenn man dem durch $w_1, w_2, \ldots$ definierten Punkt $\bar{y}$ von $M_{\mathfrak{U}}$ den durch $hw_1, hw_2, \ldots$ definierten Punkt $\bar{y}'$ von $M_{\mathfrak{U}'}$ zuordnet, so wird daher $M_{\mathfrak{U}}$ eineindeutig auf $M_{\mathfrak{U}'}$ abgebildet, und zwar so, daß $\varphi(\bar{y}) = \varphi'(\bar{y}')$ ist, wobei φ' ebenso für $M_{\mathfrak{U}'}$ definiert ist wie φ für $M_{\mathfrak{U}}$. Mithin ist wegen der Stetigkeit von φ und der eindeutigen und stetigen Umkehrbarkeit von φ' im Kleinen die eineindeutige Beziehung zwischen $M_{\mathfrak{U}}$ und $M_{\mathfrak{U}'}$ auch stetig. Damit ist bewiesen:

Zu ähnlichen Untergruppen von $\mathfrak{F}$ gehört — bis auf Homöomorphismen — dieselbe Überlagerungsmannigfaltigkeit.

§ 2.

Definition von Klasseninvarianten und deren Haupteigenschaften.

1. Der Gruppenhomomorphismus der Abbildung. M und μ seien Mannigfaltigkeiten, $\mathfrak{F}$ und Φ ihre Fundamentalgruppen, f sei eine eindeutige und stetige Abbildung von M auf μ, x sei ein Punkt von M, $\xi = f(x)$ sein Bild. Die geschlossenen Wege durch x werden durch f auf geschlossene Wege — nicht notwendigerweise auf alle geschlossenen Wege — durch ξ abgebildet, jeder zusammenziehbare Weg wird auf einen zusammenziehbaren Weg, die Zusammensetzung zweier Wege wird auf die Zusammensetzung der Bilder der beiden Wege abgebildet. Daraus folgt, daß f die Gruppe $\mathfrak{F}_x$ der geschlossenen Wege durch x homomorph in die Gruppe Φ_ξ der geschlossenen Wege durch ξ abbildet. Bei Auszeichnung von x vermittelt also f einen Homomorphismus von $\mathfrak{F}$ in Φ [14]).

Ist y ein von x verschiedener Punkt in M, $\eta = f(y)$ sein Bild, so bildet f ebenso die Gruppe $\mathfrak{F}_y$ homomorph in die Gruppe Φ_η ab. Nun sind, wenn $\mathfrak{F}_x$ und Φ_ξ in bestimmter Weise als Realisationen der Fundamentalgruppen $\mathfrak{F}$ bzw. Φ aufgefaßt werden, $\mathfrak{F}_y$ und Φ_η dadurch noch nicht in eindeutiger Weise als Realisationen der Fundamentalgruppen bestimmt, vielmehr sind (vgl. § 1, 1. Absatz) willkürliche Wege v und w von x nach y bzw. von ξ nach η

[14]) Bei einem Homomorphismus von $\mathfrak{F}$ „in" Φ braucht nicht jedes Element von Φ Bild eines Elementes von $\mathfrak{F}$ zu sein; bei einem Homomorphismus von $\mathfrak{F}$ „auf" Φ soll dies immer der Fall sein (diese Ausdrucksweise kenne ich durch Herrn van der Waerden).

einzuführen, die zwischen $\mathfrak{F}_x$ und $\mathfrak{F}_y$ bzw. zwischen Φ_ξ und Φ_η vermitteln: dem Element g von $\mathfrak{F}_x$ wird das Element $v^{-1}gv$ von $\mathfrak{F}_y$, dem Element γ von Φ_ξ wird das Element $w^{-1}\gamma w$ von Φ_η zugeordnet. Ist der Homomorphismus von $\mathfrak{F}_x$ in Φ_ξ durch Gleichungen $f(g)=\gamma$ (wobei g die Gruppe $\mathfrak{F}_x$ durchläuft) gegeben, so ist der Homomorphismus von $\mathfrak{F}_y$ in Φ_η nicht durch $f(v^{-1}gv)=w^{-1}\gamma w$, sondern durch $f(v^{-1}gv)=f(v)^{-1}f(g)f(v)$ $=(f(v)^{-1}w)w^{-1}\gamma w(f(v)^{-1}w)^{-1}$ gegeben; dabei ist $f(v)^{-1}w$ ein Element von Φ_η. Der durch f zwischen $\mathfrak{F}$ und Φ vermittelte Homomorphismus ist also — bei Auszeichnung verschiedener Punkte $x, y, \ldots$ — nur bis auf innere Automorphismen von Φ bestimmt.

f werde nun stetig abgeändert, d. h. es existiere eine von t stetig abhängende, für $0\leq t\leq 1$ erklärte Schar von Abbildungen f_t mit $f_0=f$. Es sei g ein geschlossener Weg durch x, $f_t(g)=\gamma_t$, w_τ der Weg, den das Bild von x durchläuft, während t von τ bis 1 wächst. $\gamma_1^{-1}w_\tau^{-1}\gamma_\tau w_\tau$ ist ein geschlossener Weg durch $\xi_1=f_1(x)$; läßt man τ von 0 bis 1 wachsen, so geht der Weg $\gamma_1^{-1}w_0^{-1}\gamma_0 w_0$ stetig in den Weg $\gamma_1^{-1}\gamma_1$ über; dieser ist zusammenziehbar; also ist es auch $\gamma_1^{-1}w_0^{-1}\gamma_0 w_0$; mithin ist $\gamma_1\equiv w_0^{-1}\gamma_0 w_0$, $f_1(g)\equiv w_0^{-1}f_0(g)w_0$. Dies gilt — bei festem Weg w_0 — für alle Elemente g von $\mathfrak{F}_x$. Der — immer nur bis auf innere Automorphismen von Φ bestimmte — Homomorphismus von $\mathfrak{F}$ in Φ ändert sich also bei dem stetigen Übergang von f_0 zu f_1 nicht.

Somit werden wir, in dem Bestreben, Eigenschaften von f zu untersuchen, die bei stetiger Abänderung von f invariant sind, unser Augenmerk auf Eigenschaften eines Homomorphismus von $\mathfrak{F}$ in Φ richten, die sich bei inneren Automorphismen von Φ nicht ändern. Nun bilden diejenigen Elemente von Φ, die bei einem Homomorphismus H von $\mathfrak{F}$ in Φ Bildelemente sind, eine Untergruppe $H(\mathfrak{F})=\mathfrak{U}$ von Φ. Wird Φ einem inneren Automorphismus unterworfen, so ist $\mathfrak{U}$ durch eine ähnliche Untergruppe $\mathfrak{U}'=h\mathfrak{U}h^{-1}$ zu ersetzen. In unserem Falle ist, wenn wir $\mathfrak{F}$ und Φ durch $\mathfrak{F}_x$ bzw. Φ_ξ realisieren, $\mathfrak{U}$ die Gruppe derjenigen Klassen geschlossener Wege durch ξ, welche Bilder geschlossener Wege durch x enthalten. Unter den Eigenschaften, die zugleich mit $\mathfrak{U}$ den mit $\mathfrak{U}$ ähnlichen Untergruppen zukommen, die also invariant gegenüber stetiger Abänderung von f sind, betrachten wir im folgenden den Index von $\mathfrak{U}$ in Φ und die zu $\mathfrak{U}$ gehörige Überlagerungsmannigfaltigkeit von μ:

Definition I. Unter der zu der Abbildung f von M auf μ gehörigen „Bildgruppe“ verstehen wir die — bis auf innere Automorphismen von Φ bestimmte — Untergruppe $\mathfrak{U}$ von Φ, die durch die Bilder der geschlossenen Wege in M definiert wird.

Definition II. Unter dem „Index“ j von f verstehen wir den Index der Bildgruppe in Φ.

Definition III. Unter der zu f gehörigen „Überlagerungsmannigfaltigkeit“ verstehen wir die zu der Bildgruppe gehörige Überlagerungsmannigfaltigkeit μ^* von μ.

Die drei so definierten Begriffe sind, wie oben gezeigt wurde, invariant gegenüber stetiger Abänderung von f.

2. Die Zerlegung der Abbildung. z ist im folgenden stets ein fester Punkt in M, $\zeta = f(z)$ sein Bild. Φ_ζ ist die Gruppe der Klassen geschlossener Wege durch ζ, $\mathfrak{U}_\zeta$ die $\mathfrak{U}$ entsprechende Untergruppe von Φ_ζ, also die Gruppe derjenigen Wegeklassen durch ζ, die Bilder geschlossener Wege durch z enthalten. μ^* denken wir uns durch von ζ ausgehende Wege realisiert, d. h. zwei solche Wege v_1, v_2 stellen dann und nur dann denselben Punkt von μ^* dar, wenn sie mod $\mathfrak{U}_\zeta$ äquivalent sind, wenn also $v_1 v_2^{-1}$ dem Bild eines geschlossenen Weges durch z äquivalent ist.

Sind w_1, w_2 zwei beliebige Wege von z nach x, so sind $f(w_1)$, $f(w_2)$ von ζ nach $\xi = f(x)$ führende Wege, und es ist $f(w_1) f(w_2)^{-1} = f(w_1 w_2^{-1})$, also sind, da $w_1 w_2^{-1}$ ein geschlossener Weg durch z ist, $f(w_1)$ und $f(w_2)$ mod $\mathfrak{U}_\zeta$ äquivalent und stellen denselben Punkt ξ^* von μ^* dar. Diesen ordnen wir dem Punkt x zu und nennen ihn $f^*(x)$. Bezeichnet wieder φ die durch die Überlagerung bewirkte Abbildung von μ^* auf μ, so ist $\varphi(\xi^*) = \varphi f^*(x) = \xi$. Aus der im Kleinen eindeutigen und stetigen Umkehrbarkeit von φ folgt die Stetigkeit von f^*.

Damit haben wir f in zwei nacheinander auszuführende eindeutige und stetige Abbildungen zerlegt: zuerst wird M durch f^* auf μ^*, dann μ^* durch φ auf μ abgebildet; es ist $f(M) = \varphi f^*(M)$. Dabei ist φ die im § 1 ausführlich besprochene Abbildung der j-schichtigen Überlagerungsmannigfaltigkeit μ^* von μ auf μ, und j ist der Index von f. f^* hat den Index 1; denn ist γ^* ein geschlossener Weg durch denjenigen Punkt ζ^* von μ^*, der den von ζ nach ζ führenden, zu $\mathfrak{U}_\zeta$ gehörigen Wegen entspricht, so gehört $\gamma = \varphi(\gamma^*)$ zu $\mathfrak{U}_\zeta$ (da andernfalls der in ζ^* beginnende Weg γ^* in einem anderen Punkte enden müßte), es gibt also einen geschlossenen Weg g durch z, so daß $\gamma f(g)^{-1}$ zusammenziehbar ist; nun ist $\gamma f(g)^{-1} = \varphi(\gamma^* f^*(g)^{-1})$, und daraus folgt (vgl. § 1), daß $\gamma^* f^*(g)^{-1}$ auf μ^* zusammenziehbar, daß also γ^* dem Bild eines geschlossenen Weges äquivalent ist. Bei dem durch f^* bewirkten Homomorphismus von $\mathfrak{F}$ in die Fundamentalgruppe $\mathfrak{U}$ von μ^* ist also jedes Element von $\mathfrak{U}$ Bild, d. h. der Index ist 1.

Wird f stetig abgeändert, d. h. liegt eine von t stetig abhängende Schar f_t von Abbildungen mit $f = f_0$ vor, so ändert sich auch $f_t(z) = \zeta_t$ stetig mit t, und μ^* ist für verschiedene Werte von t durch von verschiedenen Punkten ζ_t ausgehende Wege zu realisieren; die homöomorphe

Beziehung zwischen diesen verschiedenen Realisationen μ_t^* ist (siehe § 1) noch abhängig von der Wahl willkürlicher Wege zwischen den verschiedenen Punkten ζ_t. Wir beseitigen diese Unbestimmtheit durch die Festsetzung: v_τ sei die während des Intervalls $0 \leqq t \leqq \tau$ von ζ_t durchlaufene Bahn; der durch den von ζ_τ ausgehenden Weg u dargestellte Punkt von μ_τ^* wird identifiziert mit dem Punkt von μ_0^*, der durch den Weg $v_\tau u$ dargestellt wird. Hierdurch wird (siehe § 1) die unter Zugrundelegung der Gruppe $\mathfrak{U}_\tau$ der geschlossenen Bildwege $f_\tau(g)$ konstruierte Überlagerungsmannigfaltigkeit μ_τ^* mit derjenigen zu ζ_0 gehörigen Überlagerungsmannigfaltigkeit identifiziert, deren Konstruktion die Gruppe $v_\tau \mathfrak{U}_\tau v_\tau^{-1}$ zugrunde gelegt ist; diese Untergruppe ist aber mit der Gruppe $\mathfrak{U}_0 = \mathfrak{U}_\zeta$ der Bilder $f_0(g)$ identisch, was man erkennt, wenn t stetig von τ nach 0 läuft; die zugehörige Überlagerungsmannigfaltigkeit ist also μ_0^*. Fassen wir auf Grund dieser Beziehung zwischen μ_t^* und μ_0^* die zu f_t gehörige Abbildung f_t^* als Abbildung auf $\mu_0^* = \mu^*$ auf, so ist das Bild eines Punktes x von M folgendermaßen zu bestimmen: man nimmt einen festen Weg w von z nach x und ordnet x als Bild denjenigen Punkt von $\mu_0^* = \mu^*$ zu, der zu dem in ζ beginnenden Weg $v_t f_t(w)$ gehört. Dieser Bildpunkt $f_t^*(x)$ ändert sich offenbar stetig mit t, einer stetigen Änderung von f entspricht also eine stetige Änderung von f^*.

Es liegt also folgender Sachverhalt vor:

Satz I. *f läßt sich in zwei Abbildungen f^* und φ zerlegen, $f(M) = \varphi f^*(M)$; dabei ist f^* eine Abbildung vom Index 1 auf die zu f gehörige Überlagerungsmannigfaltigkeit μ^* und ändert sich bei stetiger Änderung von f stetig; φ ist die durch die Überlagerung gegebene Abbildung von μ^* auf μ und bleibt bei stetiger Änderung von f fest.*

3. Die Einteilung der Originalmenge eines Bildpunktes in Schichten. ξ sei ein Punkt in μ, X seine Originalmenge in M, d. h. die Menge derjenigen Punkte x, für die $f(x) = \xi$ ist. Ist u ein Weg zwischen zwei Punkten x_1, x_2 von X, so ist $f(u)$ ein geschlossener Weg in μ. Es kann eintreten, daß dieser zusammenziehbar ist; dies sei der Fall, und es sei ferner v ein Weg von x_2 nach dem ebenfalls zu X gehörigen Punkt x_3: dessen Bild $f(v)$ auch zusammenziehbar ist. Dann ist das Bild $f(uv) = f(u)f(v)$ des von x_1 nach x_3 laufenden Weges uv ebenfalls zusammenziehbar. Diese Tatsache ermöglicht es, die Menge X in zueinander fremde Teilmengen X_1, $X_2, \ldots$ durch die Bestimmung einzuteilen, daß zwei Punkte von X dann und nur dann derselben Teilmenge X_i angehören, wenn sie sich durch einen Weg verbinden lassen, dessen Bild zusammenziehbar ist; diese Teilmengen X_i bezeichnen wir als „Schichten", d. h. wir definieren:

Definition IV. Eine Teilmenge der Originalmenge X von ξ heißt „Schicht", wenn sich je zwei ihrer Punkte durch einen Weg verbinden

lassen, dessen Bild, als geschlossener Weg in μ aufgefaßt, zusammenziehbar ist, und wenn sie nicht in einer größeren Teilmenge von X enthalten ist, die dieselbe Eigenschaft hat.

Ist x ein Häufungspunkt von X, so gehört er wegen der, aus der Stetigkeit von f folgenden, Abgeschlossenheit von X zu X und somit zu einer bestimmten Schicht, etwa zu X_1. Ist U eine so kleine Umgebung von x, daß ihr Bild in einem ξ enthaltenden euklidischen Element E enthalten ist, und sind x_1, x_2 irgend zwei Punkte von X in U, so ist das Bild eines Weges u, der x_1 mit x_2 innerhalb U verbindet, ein geschlossener Weg in E, also zusammenziehbar; mithin gehören x_1, x_2 zu einer Schicht, und zwar, da $x_1 = x$ sein kann, zu X_1. Ein Häufungspunkt von X ist also stets Häufungspunkt einer und nur einer Schicht und gehört dieser an. Hierin ist erstens enthalten, daß jede Schicht abgeschlossen ist; bezeichnen wir ein System von Mengen $A_1, A_2, \ldots$ in M als „isoliert", wenn es um jeden Punkt von M eine Umgebung gibt, welche Punkte höchstens einer der Mengen A_i enthält, so haben wir zweitens gezeigt, daß die Schichten $X_1, X_2, \ldots$ ein isoliertes System bilden; es gilt also:

Satz II. *Die zu einem Punkt ξ gehörigen Schichten bilden ein isoliertes System abgeschlossener Mengen.*

Besteht ein isoliertes System $A_1, A_2, \ldots$ aus unendlich vielen, nicht leeren Mengen A_i, so kann eine unendliche Punktmenge $a_1, a_2, \ldots$, die aus Punkten $a_i \subset A_i$ gebildet ist, wegen der Eigenschaft der Isoliertheit keinen Häufungspunkt in M haben, die Vereinigungsmenge der A_i ist also nicht kompakt. Damit ist gezeigt:

Satz IIa. *Wenn die Originalmenge X von ξ kompakt ist — insbesondere also, wenn die Abbildung f im Punkte ξ kompakt ist*[15]), — *so besteht X nur aus endlich vielen Schichten.*

4. Der Zusammenhang zwischen der Zerlegung von f und der Schichteneinteilung. Neben der Einteilung von X in Schichten läßt sich X auf Grund der in Nr. 2 besprochenen Zerlegung von f in f^* und φ dadurch in zueinander fremde Teile zerspalten, daß man Punkte von X dann und nur dann zum selben Teil rechnet, wenn sie durch f^* auf denselben Punkt von μ^* abgebildet werden. Diese beiden Einteilungen von X sind aber miteinander identisch, d. h. es gilt

Satz III. *Zwei Punkte x_1, x_2 von X gehören dann und nur dann zu derselben Schicht, wenn $f^*(x_1) = f^*(x_2)$ ist.*

Beweis. x_1 und x_2 mögen zu derselben Schicht gehören. Dann gibt es einen Weg u von x_1 nach x_2 derart, daß der durch $\xi = f(x_1) = f(x_2)$

[15]) Definition der „Kompaktheit von f in ξ": Teil I, § 4, S. 598.

laufende geschlossene Weg $f(u)$ zusammenziehbar ist. w_1 sei ein beliebiger Weg von z (siehe Nr. 2) nach x_1, w_2 der durch $w_2 = w_1 u$ definierte Weg von z nach x_2. Dann ist $f(w_1)f(w_2)^{-1} = f(w_1)f(w_1 u)^{-1} = f(w_1)f(u)^{-1}f(w_1)^{-1}$; dabei läuft $f(w_1)$ von $\zeta = f(z)$ nach ξ, $f(u)^{-1}$ von ξ nach ξ zurück, $f(w_1)^{-1}$ von ξ nach ζ. $f(u)^{-1}$ läßt sich unter Festhaltung von ξ auf ξ zusammenziehen; dadurch geht $f(w_1)f(w_2)^{-1}$ in den Weg $f(w_1)f(w_1)^{-1}$ über, der selbst zusammenziehbar ist; mithin ist $f(w_1)f(w_2)^{-1}$ zusammenziehbar, $f(w_1)$ und $f(w_2)$ sind also gewiß äquivalent mod $\mathfrak{U}_\zeta$, d. h. es ist $f^*(x_1) = f^*(x_2)$.

Es sei andererseits $f^*(x_1) = f^*(x_2)$. Dann ist, wenn w_1, w_2 Wege von z nach x_1, x_2 sind, $f(w_1)f(w_2)^{-1} \subset \mathfrak{U}_\zeta$, es gibt also einen geschlossenen Weg g durch z, so daß $f(w_1)f(w_2)^{-1} = f(g)$ ist; das bedeutet, daß $f(w_1)f(w_2)^{-1}f(g)^{-1}$ zusammenziehbar ist; setzen wir $w_2^{-1}g^{-1} = u$, so ist u ein Weg von x_2 nach z, und $f(w_1)f(u)$ ist zusammenziehbar. Nun folgt aber stets, wenn a und b zwei Wege sind, derart, daß der Anfangspunkt des einen mit dem Endpunkt des anderen zusammenfällt, aus der Zusammenziehbarkeit von ab die Zusammenziehbarkeit von ba. Daher ist auch $f(u)f(w_1) = f(uw_1)$ zusammenziehbar, und da uw_1 ein Weg von x_2 nach x_1 ist, ist gezeigt, daß x_1 und x_2 zu derselben Schicht gehören.

Damit ist der Satz III bewiesen.

Da bei der Abbildung φ der Punkt ξ genau j Originalpunkte auf μ^* hat, folgt aus den Sätzen I und III

Satz IIIa. *Die Anzahl der zu einem Punkt ξ gehörigen Schichten ist höchstens gleich dem Index j von f*[16]).

5. Orientierbarkeit einer Abbildung. Sind x, y Punkte in M, E_x, E_y Elemente, die diese Punkte enthalten, und ist u ein Weg von x nach y, so läßt sich eine in E_x willkürlich ausgezeichnete Orientierung folgendermaßen längs u auf E_y übertragen: Durch Einfügung von Punkten $x_1, x_2, \ldots, x_{k-1}$ teilt man u in Teilwege $u_1 = x x_1, \ldots, u_i = x_{i-1} x_i, \ldots, u_k = x_{k-1} y$ ein, die so klein sind, daß sich jeder Bogen u_i in ein Element E_i einschließen läßt, wobei $E_1 = E_x$, $E_k = E_y$ ist; man überträgt nacheinander für $i = 1, 2, \ldots, k-1$ die Orientierung von E_i vermöge der Koinzidenz in einer Umgebung von x_i auf E_{i+1} und gelangt somit zu einer bestimmten Orientierung von E_y. Diese Orientierung hängt weder von der Wahl der

[16]) Bis hierher ist von den Räumen M und μ nicht benutzt worden, daß sie Mannigfaltigkeiten, sondern nur, daß sie zusammenhängende, im Kleinen zusammenhängende, im Kleinen kompakte topologische Räume sind, in denen sich jeder hinreichend kleine geschlossene Weg zusammenziehen läßt. — Von nun an aber ist es wesentlich, daß M und μ Mannigfaltigkeiten von der gleichen Dimensionszahl n sind.

Punkte x_i noch von der Wahl der Elemente E_i, also — nach Auszeichnung der Orientierung in E_x — nur von dem Wege u ab; denn man erkennt mühelos, daß sie sich erstens nicht ändert, wenn man unter Festhaltung der x_i die E_i $(i = 2, 3, \ldots, k-1)$ durch andere Elemente E_i' ersetzt, daß sie sich zweitens nicht ändert, wenn man die durch die x_i bewirkte Zerlegung von u in die Bögen u_i durch Einfügung neuer Teilpunkte verfeinert, und daß sie schließlich von der Wahl der x_i unabhängig ist, da man zu zwei verschiedenen Unterteilungen von u stets eine beiden gemeinsame feinere Unterteilung angeben kann.

Diese Übertragung der Orientierung von E_x auf E_y längs u behält ihren Sinn, wenn x mit y zusammenfällt, wenn also u geschlossen ist. Es sind dann zwei Fälle möglich: die vermöge der Koinzidenz in x von E_y auf E_x übertragene Orientierung ist entweder mit der ursprünglichen Orientierung identisch oder sie ist ihr entgegengesetzt. Je nachdem ob der erste oder der zweite Fall eintritt sagen wir, daß der geschlossene Weg u die Orientierung erhält oder umkehrt. Man sieht übrigens leicht, daß diese Unterscheidung nicht von der Wahl des Punktes x auf u, sondern nur von dem Weg u abhängt.

Ist u' ein zu u hinreichend benachbarter[17]) geschlossener Weg, so ergibt sich aus der Definition mittels der Elemente E_i wieder ohne weiteres, daß u' die Orientierung erhält oder umkehrt, je nachdem u sie erhält oder umkehrt. Da ferner jeder ganz in einem Element verlaufende geschlossene Weg die Orientierung erhält und sich jeder zusammenziehbare Weg in das Innere eines Elementes hinein stetig zusammenziehen läßt, folgt hieraus, daß jeder zusammenziehbare Weg die Orientierung erhält. Weiter ergibt sich aus der Definition, daß die Zusammensetzung uv zweier geschlossener Wege u und v durch x die Orientierung erhält, wenn beide Wege sie erhalten oder beide sie umkehren, daß dagegen uv die Orientierung umkehrt, wenn von den beiden Wegen einer sie erhält, der andere sie umkehrt. Aus alledem folgt, daß die Eigenschaft, die Orientierung zu erhalten oder sie umzukehren, nicht nur einzelnen geschlossenen Wegen durch x, sondern den ganzen Wegeklassen zukommt, daß die Klassen, die die Orientierung erhalten, eine Untergruppe $\mathfrak{G}$ der Fundamentalgruppe bilden, daß, wenn i ein bestimmter die Orientierung umkehrender Weg durch x ist, es zu jedem die Orientierung umkehrenden Weg i' durch x einen die Orientierung erhaltenden Weg g so gibt, daß $i' = ig$ ist, daß also $\mathfrak{G}$ genau zwei Restklassen in $\mathfrak{F}$ hat und man mithin $\mathfrak{F}$ in $\mathfrak{F} = \mathfrak{G} + i\mathfrak{G}$ zerlegen kann, vorausgesetzt, daß überhaupt ein i existiert.

[17]) Dabei sind u und u' als Bilder einer festen Kreislinie k aufzufassen und die Entfernung zwischen u und u' ist das Maximum der Entfernungen der beiden Bilder eines k durchlaufenden Punktes.

Aus der Definition der Orientierbarkeit einer Mannigfaltigkeit[18]) folgt, daß kein i existiert, wenn M orientierbar ist. Ist dagegen M nicht orientierbar, so gibt es Wege i; denn andernfalls wäre die Fortsetzung der Orientierung von E_x nach einem beliebigen anderen Element E_y vom Wege unabhängig (da, wenn u, v zwei Wege von x nach y sind, der Weg $u v^{-1}$ die Orientierung erhielte), also in eindeutiger Weise möglich, M wäre also orientierbar.

Wir teilen nun die Abbildungen von M auf μ nach dem Verhalten der Bilder der die Orientierung umkehrenden geschlossenen Wege in zwei Kategorien ein:

Definition V. f heißt „nicht orientierbar", wenn es einen die Orientierung umkehrenden geschlossenen Weg in M gibt, dessen Bild in μ zusammenziehbar ist; andernfalls heißt f orientierbar.

Ist M orientierbar, so ist jede Abbildung f orientierbar.

Ist M nicht orientierbar, f orientierbar und ist g ein die Orientierung erhaltender, i ein die Orientierung umkehrender geschlossener Weg durch x, so können die durch $\xi = f(x)$ laufenden geschlossenen Wege $f(g), f(i)$ nicht untereinander äquivalent sein, da sonst $f(g) f(i)^{-1} = f(g i^{-1})$ zusammenziehbar wäre, was der Orientierbarkeit von f widerspräche, da $g i^{-1}$ ein geschlossener Weg ist, der die Orientierung umkehrt. Also ist ein geschlossener Weg in μ, sofern er überhaupt dem Bilde eines geschlossenen Weges von M äquivalent ist, entweder nur Bildern von Wegen, die die Orientierung erhalten, oder nur den Bildern von Wegen, die die Orientierung umkehren, äquivalent.

Ist f nicht orientierbar, i ein die Orientierung umkehrender Weg, für den $f(i)$ zusammenziehbar ist, z irgendein Punkt in M, w ein Weg von z nach einem Punkt x von i, so kehrt der durch z gehende geschlossene Weg $i' = w i w^{-1}$ die Orientierung um und sein Bild ist ebenfalls zusammenziehbar; derartige geschlossene Wege gibt es also durch jeden Punkt von M. Ist g irgendein die Orientierung erhaltender oder umkehrender, geschlossener Weg durch z, so kehrt $i' g$ die Orientierung um bzw. erhält sie; die Bilder $f(g)$ und $f(i' g)$ sind äquivalent, da $f(i')$ zusammenziehbar ist. Also ist ein geschlossener Weg in μ, sofern er überhaupt dem Bilde eines geschlossenen Weges von M äquivalent ist, sowohl dem Bild eines die Orientierung erhaltenden, als dem Bild eines die Orientierung umkehrenden Weges äquivalent.

Ist f orientierbar, X_i eine Schicht, x_1 ein Punkt von X_i, in dessen Umgebung eine Orientierung ausgezeichnet ist, so ist, wenn man für die Fortsetzung dieser Orientierung nach anderen Punkten von X_i nur solche

[18]) Teil I, § 3.

Wege zuläßt, deren Bilder in μ zusammenziehbar sind, die so fortgesetzte Orientierung in der Umgebung jedes Punktes von X_i eindeutig bestimmt; denn ist x_2 ein zweiter Punkt von X_i und sind u, v zwei Wege von x_1 nach x_2, deren Bilder zusammenziehbar sind, so ist auch das Bild von $u v^{-1}$ zusammenziehbar, also erhält wegen der Orientierbarkeit von f der Weg $u v^{-1}$ die Orientierung, mithin liefert die Fortsetzung der Orientierung der Umgebung von x_1 längs u dieselbe Orientierung der Umgebung von x_2 wie die Fortsetzung längs v; zu jeder der beiden Orientierungen der Umgebung von x_1 gehört also eine bestimmte Orientierung einer Umgebung der ganzen Schicht X_i, und wir dürfen festsetzen:

Definition VI. Bei einer orientierbaren Abbildung f verstehen wir unter einer Orientierung der Umgebung einer Schicht X_i solche Orientierungen von euklidischen Umgebungen der Punkte von X_i, die aus einer Orientierung der Umgebung eines Punktes von X_i durch Fortsetzung längs Wegen hervorgehen, deren Bilder zusammenziehbar sind.

Ist M orientierbar, so ist eine solche Orientierung natürlich die durch eine der Orientierungen von ganz M bewirkte.

6. **Die Beiträge der Schichten.** f sei im Punkte ξ von μ kompakt. G sei eine offene, die Originalmenge X von ξ enthaltende Menge in M und zerfalle in die Komponenten $G_1, G_2, \ldots$; dabei sind die G_i zueinander fremde Gebiete, d. h. zusammenhängende offene Mengen in M, also Teilmannigfaltigkeiten von M; und zwar sind sie sämtlich, wenn nicht $G = M$ und M geschlossen ist, offene Mannigfaltigkeiten.

Nur in endlich vielen G_i können Punkte von X enthalten sein; denn würde jedes Gebiet einer unendlichen Folge $G_1, G_2, \ldots$ von Komponenten je einen Punkt $x_1, x_2, \ldots$ von X enthalten, so hätte die Folge der x_i wegen der Kompaktheit von X einen Häufungspunkt x_0, dieser gehörte einer Komponente G_0 an und mit dieser hätten fast alle der Gebiete $G_1, G_2, \ldots$ Punkte — nämlich Punkte von X — gemeinsam, im Widerspruch zu ihrer Eigenschaft, voneinander verschiedene Komponenten zu sein.

Die endlich vielen Mannigfaltigkeiten G_i, die Punkte von X enthalten, mögen die „wesentlichen“ Komponenten von G heißen.

Ist G_i eine wesentliche Komponente, so ist die Abbildung $f(G_i)$ in ξ kompakt (für eine unwesentliche Komponente ist dies trivial). Ist nämlich die positive Zahl δ erstens kleiner als der Abstand des Punktes ξ von dem Bilde des Randes von G_i und zweitens so klein, daß die ganze Originalmenge der δ-Umgebung von ξ in M kompakt ist — infolge der Kompaktheit von $f(M)$ in ξ ist diese Bedingung erfüllbar —, so hat der zu G_i gehörende Teil dieser Originalmenge keinen Häufungspunkt auf dem Rande von G_i, ist also in G_i kompakt.

Wir nehmen als G nun die ε-Umgebung von X, wobei ε eine die folgenden zwei Bedingungen erfüllende positive Zahl ist: erstens ist die Bildmenge $f(G)$ in einem ξ enthaltenden euklidischen Element enthalten; zweitens ist 2ε kleiner als der kleinste Abstand zwischen irgend zwei von den nach Satz IIa in endlicher Anzahl vorhandenen Schichten, in die X zerfällt. Dann enthält infolge der zweiten Bedingung keine der Mannigfaltigkeiten G_i Punkte zweier verschiedener Schichten. Jede der wesentlichen Komponenten G_i enthält also Punkte genau einer Schicht; und zwar seien diejenigen wesentlichen Komponenten von G, die Punkte von X_i enthalten, jetzt mit $G_{i1}, G_{i2}, \ldots$ bezeichnet.

Ist f orientierbar, so ist jede Mannigfaltigkeit G_{ij} orientierbar; denn ist g ein geschlossener Weg in G_{ij}, so ist infolge der ersten ε auferlegten Bedingung $f(g)$ in einem euklidischen Element enthalten, also zusammenziehbar, mithin erhält wegen der Orientierbarkeit von f der Weg g die Orientierung. Die Orientierung von G_{i1} bewirkt auf Grund von Definition VI eine bestimmte Orientierung von $G_{i2}, G_{i3}, \ldots$.

Die Abbildung jeder der so orientierten Mannigfaltigkeiten G_{ij} (i fest, $j = 1, 2, \ldots$) hat, da sie, wie oben gezeigt wurde, in ξ kompakt ist, in ξ einen bestimmten Grad[19]), wenn man noch eine bestimmte Orientierung der Umgebung von ξ ausgezeichnet hat. Ändert man diese Orientierung oder die Orientierung von G_{i1} und damit die Orientierung von $G_{i2}, G_{i3}, \ldots$, so ändert die Summe der Grade der Abbildungen von $G_{i1}, G_{i2}, \ldots$ ihr Vorzeichen, ihr absoluter Betrag bleibt derselbe. Dieser bleibt ferner ungeändert, wenn man ε verkleinert, da dann von den G_{ij} nur Punkte fortfallen, deren Bilder ξ nicht bedecken, die also auf den Wert des Grades keinen Einfluß haben. Der absolute Betrag der Summe der Grade hängt also nur von der Schicht X_i selbst ab und werde als der „Beitrag" von X_i bezeichnet.

Ist f nicht orientierbar, so wird sowohl die Orientierbarkeit der Mannigfaltigkeiten G_{ij} in Frage gestellt, als auch die Übertragung der Orientierung von G_{i1} auf $G_{i2}, G_{i3}, \ldots$ unmöglich. Die Grade sind daher nicht definiert, die Betrachtung behält aber ihren Sinn, wenn wir die Grade durch die Paritäten[19]) ersetzen und sagen, daß die Schicht den „Beitrag" 0 oder 1 liefere, je nachdem die Summe der Paritäten der Abbildungen von $G_{i1}, G_{i2}, \ldots$ in ξ gerade oder ungerade ist.

Wir fassen die Definition der „Beiträge" noch einmal zusammen und fügen einige weitere Definitionen hinzu:

Definition VIIa. Ist f in ξ kompakt, so verstehen wir unter dem Beitrag einer zu ξ gehörigen Schicht X_i die folgende Zahl: wenn f orientier-

[19]) Teil I, § 5.

bar ist, den absoluten Wert des Grades der Abbildung einer hinreichend kleinen, gemäß Definition VI orientierten Umgebung von X_i im Punkte ξ; wenn f nicht orientierbar ist, die Zahl 0 oder 1, je nachdem die Parität der Abbildung einer hinreichend kleinen Umgebung von X_i im Punkte ξ gerade oder ungerade ist.

Definition VIIb. Eine Schicht heißt „wesentlich" oder „unwesentlich", je nachdem ihr Beitrag von 0 verschieden oder gleich 0 ist; die — nach Satz IIa a fortiori endliche — Anzahl der zu ξ gehörigen wesentlichen Schichten heißt die „wesentliche Schichtenzahl" von f in ξ und wird im folgenden mit s_ξ bezeichnet.

Definition VIIc. Die Summe der (wesentlichen) Schichtenbeiträge eines Punktes ξ heißt der „Absolutgrad" von f in ξ und wird im folgenden mit a_ξ bezeichnet.

Ist M orientierbar und sind $\gamma_1, \gamma_2, \ldots$ die auf Grund einer Orientierung von M bestimmten Grade der Umgebungen der einzelnen Schichten $X_1, X_2, \ldots$ in ξ, so ist der Grad gleich $\sum \gamma_i$, der Absolutgrad gleich $\sum |\gamma_i|$; folglich ist in diesem Fall, d. h. immer dann, wenn der Grad überhaupt definiert ist, der Absolutgrad mindestens so groß wie der absolute Betrag des Grades; insbesondere ist der Grad 0, wenn der Absolutgrad 0 ist. Ebenso folgt im Fall einer nicht orientierbaren Mannigfaltigkeit M, daß die Parität von f gerade ist, falls der Absolutgrad 0 ist.

Wir nehmen die durch Satz I gegebene Zerlegung von f vor; $\xi_1^*, \xi_2^*, \ldots$ seien die Punkte von μ^*, die durch φ auf ξ abgebildet werden. Nach Satz III gehört zu jeder Schicht X_i ein Punkt $\xi_i = f^*(X_i)$. X_i bildet auch bezüglich f^* eine einzige Schicht, da jeder durch φ auf einen zusammenziehbaren Weg von μ abgebildete Weg von μ^* selbst zusammenziehbar ist (siehe § 1). Der von X_i bezüglich f^* gelieferte Beitrag ist, da φ in der Umgebung von ξ_i^* eineindeutig und $f = \varphi f^*$ ist, gleich dem von X_i bezüglich f gelieferten Beitrag a_i; dabei hat man zu berücksichtigen, daß f^* orientierbar oder nicht orientierbar ist, je nachdem f es ist oder nicht ist, wieder infolge des Entsprechens der zusammenziehbaren Wege auf μ und μ^*.

G sei nun wieder eine ε-Umgebung von X und ε wieder so klein, daß verschiedene Schichten zu verschiedenen Komponenten von G gehören, und überdies so klein, daß Komponenten, die Punkte verschiedener Schichten X_i enthalten, durch f^* in die zueinander fremden Umgebungen U_i^* der $\xi_i^* = f^*(X_i)$ hinein abgebildet werden, die einer festen Umgebung U von ξ entsprechen. V sei eine in U enthaltene Umgebung von ξ, $V_1^*, V_2^*, \ldots$ seien die entsprechenden Umgebungen der $\xi_1^*, \xi_2^*, \ldots$; V sei so klein, daß $f(M - G)$ außerhalb V, daß also die Originalmenge von V in G liegt; diese Original-

menge besteht aus den zueinander fremden Originalmengen der V_i^* bezüglich f^*, und die Originalmenge einer einzelnen V_i^* gehört wegen $V_i^* \subset U_i^*$ solchen Komponenten von G an, die Punkte von X_i enthalten; hat ein Punkt ξ_j^* keinen Originalpunkt bei f^*, so ist auch die Originalmenge seiner Umgebung V_j^* leer.

Nun sei V ferner so klein, daß f in allen Punkten von V kompakt ist; dann ist f^* in allen Punkten von V_i^* kompakt, und da das durch f^* gelieferte Bild von $M - G$, also auch das des Randes von G fremd zu V_i^* ist, ist auch die Abbildung $f^*(G)$ und somit erst recht die Abbildung f^* jeder einzelnen Komponente von G in allen Punkten von V_i^* kompakt. Jede dieser Abbildungen hat also in V_i^* konstanten Grad bzw. konstante Parität. Ist η ein Punkt von V, η_i^* der entsprechende Punkt in V_i^*, a_i der Beitrag von X_i, so ist also a_i die Grad- bzw. Paritätensumme der Abbildung der Komponenten von G nicht nur in ξ_i^*, sondern auch in η_i^*. Da es außerhalb G keine Originalpunkte von η_i^* gibt, ist a_i die Grad- bzw. Paritätensumme bei der Abbildung der Komponenten einer Umgebung der ganzen Originalmenge Y_i von η_i^*. Ist $a_i \neq 0$, X_i also wesentlich, so ist Y_i nicht leer, sondern eine wesentliche Schicht von η mit dem Beitrag a_i. Ist $a_i = 0$, X_i also unwesentlich, so ist Y_i entweder leer oder unwesentlich.

Somit ist die Reihe $a_1, a_2, \ldots, a_s$ der Beiträge der wesentlichen Schichten in der Umgebung V von ξ konstant, und da es um jeden Punkt eines Gebietes, in dem f kompakt ist, eine solche Umgebung gibt, folgt:

Satz IVa. *Die Reihe der Beiträge der (wesentlichen) Schichten, also insbesondere auch die wesentliche Schichtenzahl s_ξ und der Absolutgrad a_ξ, sind konstant in jedem Gebiet, in dem f kompakt ist.*

Ändert man f gleichmäßig stetig so ab, daß die Abbildung immer kompakt in ξ bleibt, so bleibt bei hinreichender Kleinheit der Änderung die Originalmenge X in G enthalten; der Änderung entspricht (siehe Satz I) eine stetige Änderung von f^*; f^* bleibt in den ξ_i^* kompakt, die Originalmenge jedes Punktes ξ_i^* bleibt in den Komponenten von G enthalten, in denen sie ursprünglich war, und die Änderung der Abbildung jeder Komponente ist gleichmäßig stetig. Die Grade bzw. Paritäten der Abbildungen dieser Komponenten ändern sich daher in ξ_i^* nicht[19]). Daraus folgt:

Satz IVb. *Die Reihe der Beiträge der (wesentlichen) Schichten, also insbesondere auch die wesentliche Schichtenzahl und der Absolutgrad, im Punkte ξ sind Invarianten der „Abbildungsklasse in bezug auf ξ"*[20]).

[20]) Teil I, § 4.

§ 3.

Überall kompakte Abbildungen.

Aus den im vorigen Paragraphen behandelten Tatsachen leiten wir in diesem Paragraphen Folgerungen für den Spezialfall her, in dem f in allen Punkten von μ kompakt ist; dieser Fall umfaßt insbesondere alle Abbildungen geschlossener Mannigfaltigkeiten M.

Satz V. *Bei einer überall kompakten Abbildung sind alle Schichtenbeiträge untereinander gleich; d. h. entweder gibt es nur unwesentliche Schichten, oder es gibt nur wesentliche Schichten und die Schichtenbeiträge in jedem Punkt sind untereinander und mit den Schichtenbeiträgen in den anderen Punkten gleich.*

Beweis. X_i, Y_k seien Schichten in den Punkten ξ, η von μ mit den Beiträgen a_i bzw. b_k; dabei darf auch $\xi = \eta$ sein. Dann sind (siehe § 2, 5.) a_i, b_k auch die Beiträge der Schichten X_i bzw. Y_k in den Punkten $\xi_i^* = f^*(X_i)$ bzw. $\eta_k^* = f^*(Y_k)$ bei der Abbildung f^*. Aus der Kompaktheit von f in allen Punkten von μ folgt die Kompaktheit von f^* in allen Punkten von μ^*; aus Satz IVa, angewandt auf f^*, ergibt sich daher $a_i = b_k$.

Satz VI. *Sind M und μ orientierbar und ist f überall kompakt, so ist der Absolutgrad gleich dem absoluten Betrag des Grades.*

Beweis. Aus der Orientierbarkeit von μ folgt die Orientierbarkeit von μ^*, da die Orientierung euklidischer Elemente von μ durch die in ihnen eindeutige Umkehrung der Abbildung φ in eindeutiger Weise auf ein vollständiges, aus euklidischen Elementen bestehendes Umgebungssystem in μ^* übertragen wird. Sind M und μ orientiert, so ist also auch μ^* orientiert, und die Abbildung f^* hat einen bestimmten Grad γ, der wegen der Kompaktheit von f^* in ganz μ^* konstant ist. Sein absoluter Betrag ist der laut Satz V konstante Schichtenbeitrag bei f^*. Infolge des oben definierten Zusammenhanges zwischen den Orientierungen von μ und μ^* hat die Abbildung φ in jedem Gebiet von μ^*, in dem sie eineindeutig ist, den Grad $+1$. Nach der Produktregel für die Grade[19]) hat daher die Abbildung $f = \varphi f^*$ der Umgebung einer Schicht X_i im Punkt ξ ebenfalls den Grad γ. Gehören zu ξ t Schichten, so ist $t\gamma$ der Grad von f, $t|\gamma|$ der Absolutgrad.

Bemerkung. Ist M orientierbar, μ nicht orientierbar, so ist der Grad auch noch definiert und stets gleich 0[19]). Der Absolutgrad kann aber von 0 verschieden sein. So hat er z. B. bei der durch die Überlagerung gegebenen Abbildung der eine projektive Ebene zweischichtig überlagernden Kugel auf die projektive Ebene den Wert 2.

Satz VII. *Hat eine überall kompakte Abbildung einen von 0 verschiedenen Absolutgrad, so ist der Index der Abbildung endlich.*

Beweis. Ist der Index j der überall kompakten Abbildung f unendlich, so hat (siehe § 1) die Überlagerungsmannigfaltigkeit μ^* unendlich viele Schichten, d. h. so gibt es auf ihr zu jedem Punkt ξ von μ unendlich viele Punkte $\xi_1^*, \xi_2^*, \ldots$ mit $\varphi(\xi_i^*) = \xi$; da andererseits nach Satz IIa die Originalmenge X von ξ in nur endlich viele Schichten zerfällt und nach Satz III jede von diesen durch f^* auf einen der Punkte $\xi_1^*, \xi_2^*, \ldots$, nach Satz I aber kein nicht zu X gehöriger Punkt durch f^* auf einen Punkt ξ_i^* abgebildet wird, gibt es Punkte ξ_i^* — und sogar unendlich viele —, die bei der Abbildung f^* nicht Bildpunkte sind. Wäre nun in einem der Punkte $\xi_1^*, \xi_2^*, \ldots$ der Schichtenbeitrag von 0 verschieden, so wäre er nach Satz V in allen Punkten von μ^* von 0 verschieden, es wären also alle ξ_i^* Bildpunkte; mithin ist der Schichtenbeitrag von f^*, der zugleich der Schichtenbeitrag von f ist, gleich 0, und folglich ist auch der Absolutgrad 0.

Satz VIIa. *Der Index einer überall kompakten Abbildung ist ein Teiler des Absolutgrades; der laut Satz* V *konstante Schichtenbeitrag hat, wenn der Absolutgrad a und der Index j ist, den Wert $\frac{a}{j}$ (dabei wird $\frac{0}{\infty} = 0$ gesetzt); ist $a \neq 0$, so ist j gleich der (wesentlichen) Schichtenzahl s.*

Beweis. Wir dürfen $a \neq 0$ annehmen. Zu jedem der Punkte $\xi_1^*, \ldots, \xi_j^*$, für die $\varphi(\xi_i^*) = \xi$ ist, gehört derselbe Schichtenbeitrag $a_i = a_1$; es ist $a = \sum_{i=1}^{j} a_i = j a_1$, also $a_1 = \frac{a}{j}$. Daraus folgt weiter $a_1 \neq 0$, also sind alle j Schichten wesentlich, und es ist $j = s$.

Folgerung. Für die Abbildbarkeit einer geschlossenen Mannigfaltigkeit M auf die Mannigfaltigkeit μ mit einem vorgeschriebenen, von 0 verschiedenen, Absolutgrad a ist notwendig, daß sich $\mathfrak{F}$ homomorph so in Φ abbilden läßt, daß der Index der Bildgruppe endlich und ein Teiler von a ist. Z. B. muß, wenn $\mathfrak{F}$ eine endliche Gruppe ist, auch Φ eine endliche Gruppe sein.

Satz VIII. *Hat eine überall kompakte Abbildung den Absolutgrad $a = 1$, so ist jeder geschlossene Weg auf μ dem Bild eines geschlossenen Weges auf M äquivalent.*

Beweis. Nach Satz VIIa hat die Abbildung den Index $j = 1$, folglich ist die Bildgruppe $\mathfrak{U}$ mit der Fundamentalgruppe Φ von μ identisch.

Satz IX. *Bei einer überall kompakten Abbildung mit von 0 verschiedenem Absolutgrad ist die Anzahl der Komponenten (d. h. zueinander fremden zusammenhängenden abgeschlossenen Mengen), in die die Originalmenge X eines beliebigen Punktes ξ von μ zerfällt, mindestens gleich dem Index j von f.*

Beweis. Jede der wesentlichen Schichten, in die X zerfällt, enthält wenigstens eine Komponente. Die Anzahl der wesentlichen Schichten ist nach Satz VIIa gleich j.

Satz X. *Gibt es bei einer überall kompakten Abbildung mit von* 0 *verschiedenem Absolutgrad einen Punkt auf* μ, *dessen Originalmenge zusammenhängend ist, so ist jeder geschlossene Weg auf* μ *dem Bilde eines geschlossenen Weges auf* M *äquivalent.*

Beweis. Die Abbildung hat nach Satz IX den Index 1; folglich ist $\mathfrak{U}$ mit Φ identisch.

§ 4.

Die Anzahl der Originalpunkte und der glatten Bedeckungen eines Bildpunktes.

f sei wieder eine beliebige Abbildung von M auf μ und im Punkte ξ von μ kompakt; mit $\mathfrak{K}_\xi$ bezeichnen wir die „Abbildungsklasse in bezug auf ξ“, zu der f gehört, d. h. die Gesamtheit aller Abbildungen, die sich durch gleichmäßig stetige Abänderung unter Wahrung der Kompaktheit im Punkte ξ aus f herstellen lassen[20]). In $\mathfrak{K}_\xi$ gibt es Abbildungen, bei denen die Originalmenge von ξ aus endlich vielen Punkten besteht; deren Anzahl ist mindestens gleich der wesentlichen Schichtenzahl s_ξ, die innerhalb $\mathfrak{K}_\xi$ invariant ist (Satz IVb). Bezeichnen wir die innerhalb $\mathfrak{K}_\xi$ erreichbare Mindestzahl von Originalpunkten des Punktes ξ mit σ_ξ, so ist also $\sigma_\xi \geqq s_\xi$.

Statt nur nach der kleinsten Zahl von Originalpunkten von ξ kann man allgemeiner fragen, welche endlichen Zahlen als Anzahlen der Originalpunkte von ξ bei einer zu $\mathfrak{K}_\xi$ gehörigen Abbildung auftreten können; diese Frage wird durch die Antwort auf die Frage nach der Größe von σ_ξ mitbeantwortet: jede endliche Zahl, die nicht kleiner als σ_ξ ist, tritt auf. Denn wenn $\bar{f}$ eine Abbildung aus $\mathfrak{K}_\xi$ mit σ_ξ Originalpunkten von ξ ist, so läßt sich folgendermaßen eine zu $\mathfrak{K}_\xi$ gehörige Abbildung $\bar{f}_1$ mit $\sigma_\xi + k$ $(k > 0)$ Originalpunkten von ξ konstruieren: y sei ein Punkt von M mit $\bar{f}(y) = \eta \neq \xi$, E ein ξ und η enthaltendes Element von μ (bezüglich der Existenz von E siehe den unten ausgesprochenen Hilfssatz Ia), e ein y enthaltendes Element von M mit $\bar{f}(e) \subset E$ und $\xi \neq \bar{f}(e)$; dann kann man $\bar{f}$ im Innern von e so abändern, daß ξ dort genau k Originalpunkte bekommt[21]), und die so abgeänderte Abbildung $\bar{f}_1$ läßt sich, da das Innere von E als euklidischer R^n aufzufassen ist, ohne Änderung am Rande von e, also in stetigem Anschluß an $\bar{f}$, durch gleichförmige Bewegungen in E aus $\bar{f}$ herstellen; die

[21]) Folgt aus Teil I, § 1, Satz X, durch Anwendung auf k zueinander fremde Teilelemente von e.

Originalmenge einer Umgebung von ξ gehört dabei in jedem Augenblick zu e, ist also kompakt, mithin gehört $\bar{f}_1$ zu $\mathfrak{K}_\xi$.

σ_ξ hängt von ξ ab; jedoch gilt ebenso wie für die im § 2 behandelten Klasseninvarianten (siehe Satz IV)

Satz XIa. *σ_ξ ist konstant, wenn ξ ein Gebiet durchläuft, in dem f kompakt ist.*

Dieser Satz ist in dem folgenden allgemeineren Satz enthalten:

Satz XI. *ξ und η seien Punkte eines Gebietes G von μ, in dem f kompakt ist; dann tritt jede Punktmenge X von M, die als Originalmenge von ξ bei einer Abbildung f_1 aus $\mathfrak{K}_\xi$ auftritt, auch als Originalmenge von η bei einer Abbildung f_2 aus $\mathfrak{K}_\eta$ auf.*

Beim Beweis werden wir, wie oben bereits einmal, den folgenden Hilfssatz verwenden, den wir zusammen mit anderen Hilfssätzen im nächsten Paragraphen beweisen werden:

Hilfssatz Ia. Zu zwei Punkten einer Mannigfaltigkeit gibt es stets ein sie im Inneren enthaltendes Element der Mannigfaltigkeit.

Beweis von Satz XI. E sei ein ξ und η im Inneren enthaltendes Element von G. D_t sei eine topologische Deformation von μ in sich, die außerhalb und auf dem Rande von E die Identität ist und, während t von 0 bis 1 läuft, ξ in η überführt; dabei ist D_0 auch im Inneren von E die Identität. Alle Abbildungen $D_t f(M)$ sind in G kompakt; also gehört insbesondere $D_1 f(M)$ zu $\mathfrak{K}_\eta$. f_τ $(0 \leqq \tau \leqq 1)$ sei die in ξ kompakte Abbildungsschar, die $f = f_0$ in f_1 überführt. Die Abbildungsschar $D_1 f_\tau(M)$ ist in η kompakt, da die Originalmenge einer Umgebung von η bei $D_1 f_\tau$ mit der Originalmenge einer Umgebung von ξ bei f_τ identisch ist; folglich ist $D_1 f_1 = f_2 \subset \mathfrak{K}_\eta$. Die Originalmenge von η bei f_2 ist mit der Originalmenge von ξ bei f_1 identisch, sie ist also die Menge X.

Es liegt nahe, der Zahl σ_ξ eine in ähnlicher Weise definierte Zahl $\bar{\sigma}_\xi$ an die Seite zu stellen: $\bar{\sigma}_\xi$ sei die durch eine Abbildung aus $\mathfrak{K}_\xi$ erreichbare Mindestzahl der *Komponenten* (d. h. untereinander fremden, zusammenhängenden abgeschlossenen Teilmengen) der Originalmenge von ξ. Auch diese Zahl ist mindestens gleich der wesentlichen Schichtenzahl s_ξ; ferner folgt aus den Definitionen unmittelbar $\bar{\sigma}_\xi \leqq \sigma_\xi$. Die neue Definition liefert aber keine neue Zahl; es ist vielmehr stets $\bar{\sigma}_\xi = \sigma_\xi$. Diese Tatsache ist in dem folgenden Satz enthalten:

Satz XII. *Besteht bei der Abbildung f die Originalmenge X von ξ aus genau k Komponenten, so läßt sich f durch eine stetige Abänderung, die sich auf eine beliebig kleine Umgebung von X beschränkt, in eine zu $\mathfrak{K}_\xi$ gehörige Abbildung f_1 überführen, bei der die Originalmenge von ξ aus genau k Punkten besteht.*

Beweis. Wir beweisen den Satz zunächst für den Spezialfall $k = 1$; X sei also zusammenhängend. ε sei eine so kleine positive Zahl, daß das durch f gelieferte Bild der ε-Umgebung von X in einer euklidischen Umgebung U von ξ enthalten ist. Diese ε-Umgebung von X kann in mehrere Komponenten zerfallen; aber da X zusammenhängend ist, enthält nur eine Komponente Punkte von X; diese Komponente heiße M'. Durch eine beliebig kleine Abänderung in einer beliebig kleinen Umgebung von X kann man unter Wahrung der Kompaktheit in ξ die Abbildung f der Mannigfaltigkeit M' so abändern, daß ξ bei dem Ergebnis f' nur endlich viele Originalpunkte hat[22]). Ist deren Anzahl größer als eins, so läßt sie sich folgendermaßen vermindern: man schließe zwei Originalpunkte in ein Element E von M' ein (s. Hilfssatz Ia). f' bewirkt eine Abbildung des Elements E in das als euklidischen R^n aufzufassende Gebiet U. Daher[23]) kann man f', ohne außerhalb und auf dem Rande von E etwas zu ändern, im Inneren von E so abändern, daß ξ dort nur noch einen einzigen Originalpunkt hat; dabei bleibt die Originalmenge einer Umgebung von ξ immer kompakt. Mithin kann man die Originalmenge von ξ innerhalb von $\mathfrak{K}_\xi$ schließlich bis auf einen einzigen Punkt vermindern. Damit ist der Spezialfall $k = 1$ erledigt.

Ist $k > 1$, so seien K_i $(i = 1, \ldots, k)$ die Komponenten von X. Man wähle ein so kleines positives δ, daß die δ-Umgebungen der verschiedenen K_i zueinander fremd sind. Unter den Komponenten der δ-Umgebung von K_i ist, da K_i zusammenhängend ist, nur eine, die Punkte von K_i enthält. Sie ist eine Mannigfaltigkeit M_i. Bei der Abbildung $f(M_i)$ ist die Originalmenge von ξ zusammenhängend. Auf Grund des bereits bewiesenen Spezialfalles läßt sich f daher innerhalb von $\mathfrak{K}_\xi$ so abändern, daß jede Komponente K_i durch einen einzigen Originalpunkt von ξ ersetzt wird und daß weitere Originalpunkte nicht hinzutreten.

Bei der Definition von σ_ξ waren wir von der Tatsache ausgegangen, daß es in $\mathfrak{K}_\xi$ „fastglatte" Abbildungen[22]) gibt, d. h. solche, bei denen die Originalmenge von ξ endlich ist. Nun gibt es in $\mathfrak{K}_\xi$ sogar Abbildungen, die „glatt" sind, d. h. die eine endliche Anzahl von Gebieten eineindeutig auf eine Umgebung von ξ abbilden, während die Komplementärmenge dieser Gebiete keinen Originalpunkt von ξ enthält[24]). Diese Tatsache führt zu der folgenden Definition: α_ξ sei die Mindestzahl von glatten (d. h. eineindeutigen) Bedeckungen einer Umgebung von ξ, die durch eine Abbildung aus $\mathfrak{K}_\xi$ geliefert wird. α_ξ ist mindestens so groß wie der Absolut-

---

[22]) Teil I, § 4, Satz Ia.

[23]) Teil I, § 1, Satz X.

[24]) Teil I, § 4, Satz II.

grad a_ξ im Punkte ξ; denn bei einer in ξ glatten Abbildung gehören zu einer Schicht, deren Beitrag a_i ist, wenigstens a_i Originalpunkte von ξ, da die eineindeutige Abbildung der Umgebung eines solchen Punktes den Grad ± 1 bzw. ungerade Parität hat. Es ist also immer $\alpha_\xi \geqq a_\xi$.

Analog wie am Anfang des Paragraphen bei der Betrachtung von σ_ξ kann man statt nach der Mindestzahl nach jeder innerhalb $\mathfrak{K}_\xi$ überhaupt möglichen Anzahl glatter Bedeckungen von ξ fragen. Auch jetzt ist die Antwort auf diese Frage gegeben, wenn man α_ξ kennt: jede endliche Zahl, die nicht kleiner als α_ξ und die mit α_ξ kongruent modulo 2 ist, kann als Zahl glatter Bedeckungen von ξ auftreten. Denn wie früher kann man, wenn eine in ξ glatte Abbildung vorliegt, diese in einem Element e von M, dessen Bild ξ nicht enthält, stetig so abändern, daß das Bild von e ξ bedeckt; dabei kann man die abgeänderte Abbildung als glatt in ξ annehmen und, da die Abbildung von e in ξ den Grad 0 hat, als Anzahl der Bedeckungen der Umgebung von ξ durch das Bild von e jede gerade Zahl willkürlich vorschreiben[25]).

Ferner gilt in Analogie zu Satz XIa

Satz XIb. *α_ξ ist konstant, wenn ξ ein Gebiet durchläuft, in dem f kompakt ist.*

Beweis. ξ und η seien Punkte des Gebietes G, in dem f kompakt ist; f_1 sei eine Abbildung aus $\mathfrak{K}_\xi$, die in ξ glatt ist, und zwar werde bei ihr eine Umgebung von ξ k-mal glatt bedeckt. Dann liefert die im Beweise von Satz XI benutzte Konstruktion eine Abbildung f_2 aus $\mathfrak{K}_\eta$, bei der eine Umgebung von η k-mal glatt bedeckt wird. Jede Zahl k, die als Zahl der Bedeckungen einer Umgebung von ξ bei einer in ξ glatten Abbildung aus $\mathfrak{K}_\xi$ auftritt, tritt also auch als Zahl der Bedeckungen einer Umgebung von η bei einer in η glatten Abbildung aus $\mathfrak{K}_\eta$ auf, und umgekehrt. Daraus folgt die Behauptung.

Die Klasseninvarianten σ_ξ und α_ξ sind durch die Abbildung f und den Punkt ξ eindeutig bestimmt; jedoch ist aus ihrer Definition nicht ersichtlich, wie sie sich ermitteln lassen, wenn f und ξ gegeben sind (hierin besteht ein prinzipieller Gegensatz zwischen σ_ξ und α_ξ einerseits und den im § 2 behandelten Klasseninvarianten andererseits). σ_ξ und α_ξ sind durch ihre Definitionen ja nicht als Eigenschaften von f, sondern als Eigenschaften der durch f und ξ bestimmten Abbildungsmenge $\mathfrak{K}_\xi$ charakterisiert. Es entsteht daher die Aufgabe, σ_ξ und α_ξ durch Größen auszudrücken, die bereits bei vollständiger Kenntnis der Abbildung f selbst bekannt sind.

Diese Aufgabe bildet den weiteren Inhalt dieser Arbeit. Sie wird zwar für die meisten Fälle, jedoch nicht in voller Allgemeinheit gelöst

---

[25]) Teil I, § 1, Satz IX.

werden; ist die Dimensionszahl von M und μ $n \neq 2$, so wird sie dadurch gelöst, daß gezeigt wird: in den uns schon bekannten Beziehungen $\sigma_\xi \geq s_\xi$, $\alpha_\xi \geq a_\xi$ stehen immer die Gleichheitszeichen; dagegen gelten die Gleichungen $\sigma_\xi = s_\xi$, $\alpha_\xi = a_\xi$ nachweislich nicht bei gewissen Flächenabbildungen, und in diesen Fällen bleibt die oben formulierte Aufgabe ungelöst.

Nach Voranschickung einiger Hilfssätze im § 5 wird im § 6 der Fall $n \neq 2$ in dem genannten Sinne erledigt werden; im § 7 wird gezeigt werden, daß der Fall $n = 2$ sich nicht nur unserer Beweismethode entzieht, sondern daß in ihm die für die anderen Dimensionszahlen gültigen Sätze tatsächlich falsch sind.

§ 5.

Hilfssätze.

1. Sind w_1, w_2 zwei Wege mit dem gemeinsamen Anfangspunkt x und dem gemeinsamen Endpunkt y in der Mannigfaltigkeit M, so nennen wir sie, in Verallgemeinerung der im § 1 für geschlossene Wege getroffenen Festsetzungen, „äquivalent", wenn man den einen unter Festhaltung von Anfangs- und Endpunkt in den anderen deformieren kann; es ist leicht zu sehen, daß dies gleichbedeutend mit der Zusammenziehbarkeit des geschlossenen Weges $w_1 w_2^{-1}$ ist. Die Wege von x nach y werden so in Äquivalenzklassen eingeteilt.

Ist t ein Teilbogen des Weges w, ist also $w = s t u$, wobei s der Teilbogen von w von dem Anfangspunkt von w bis zu dem Anfangspunkt von t, u der Teilbogen von dem Endpunkt von t bis zu dem Endpunkt von w ist, und ist t mit t' äquivalent, so wird durch die Deformation von t in t' zugleich w in $w' = s t' u$ deformiert. Ersetzt man also einen Teilbogen eines Weges durch einen äquivalenten Bogen, so bleibt die Äquivalenzklasse des ganzen Bogens ungeändert.

Ist E ein x und y enthaltendes Element, so sind zwei Wege w_1 und w_2, die x mit y innerhalb E verbinden, stets miteinander äquivalent, da man die Überführung von w_1 in w_2 mit Hilfe der euklidischen Geometrie von E geradlinig und gleichförmig durchführen kann. Zu jedem x und y enthaltenden Element E gehört also eine wohlbestimmte Äquivalenzklasse von Wegen zwischen x und y. Es fragt sich, ob man umgekehrt bei gegebenen Punkten x und y zu jeder Äquivalenzklasse ein solches Element E angeben kann. Man sieht leicht, daß diese Frage nicht in voller Allgemeinheit zu bejahen ist: denn ist $n = 1$, M ein Kreis und w ein Weg von x nach y, der den Kreis mehrere Male umläuft, so daß die Änderung des Winkelarguments auf w größer als 2π ist, so ist sie auch auf jedem zu w äquivalenten Weg $> 2\pi$; ein solcher Weg läßt sich aber nicht in ein

Intervall einschließen. Jedoch spielt der Fall $n=1$ eine Ausnahmerolle; denn es gilt:

Hilfssatz I. Ist $n>1$, M eine n-dimensionale Mannigfaltigkeit und w ein Weg zwischen den voneinander verschiedenen Punkten x und y in M, so gibt es in M ein x und y enthaltendes Element E mit der Eigenschaft, daß die x mit y innerhalb E verbindenden Wege mit w äquivalent sind.

Bemerkung. Hierin ist der im vorigem Paragraphen ausgesprochene und benutzte „Hilfssatz Ia" für $n>1$ enthalten; da dieser aber für den Fall $n=1$, in dem M entweder eine Gerade oder ein Kreis ist, sofort zu verifizieren ist, darf er nach Beweis des Hilfssatzes I als bewiesen angesehen werden.

Beweis. e sei ein x enthaltendes, y nicht enthaltendes Element. w hat Punkte mit dem Rande r von e gemeinsam, und unter diesen gibt es einen bei der Durchlaufung von w in der Richtung xy ersten Punkt z; falls der Bogen xz von w nicht doppelpunktfrei ist, ersetzen wir ihn durch einen in e von x nach z laufenden doppelpunktfreien Bogen. Hierdurch wird, da die beiden Bögen xz in e verlaufen, also, wie früher bemerkt, äquivalent sind, die Äquivalenzklasse des ganzen Weges w, wie ebenfalls früher bemerkt, nicht geändert. Falls w in seinem weiteren Verlauf wieder in e eintritt, falls es also auf dem Bogen zy einen im Inneren von e liegenden Punkt z gibt, so gehört dieser einem Teilbogen von w an, dessen Endpunkte z_1, z_2 auf dem Rande r von e liegen; da r eine $(n-1)$-dimensionale Sphäre und $n-1>0$ ist, lassen sich z_1 und z_2 auf r durch einen Bogen verbinden, der wieder mit dem zwischen z_1 und z_2 verlaufenden Teil von w äquivalent ist, durch den wir den letzteren also ersetzen dürfen, ohne die Klasse von w zu ändern. Indem wir dies für jeden in e eintretenden Bogen tun, erreichen wir, daß der ganze Weg zunächst doppelpunktfrei von x bis in den Randpunkt z von e läuft und danach nie mehr ins Innere von e eintritt, so daß insbesondere der Anfangspunkt x nicht noch ein zweites Mal von dem Wege erreicht wird. Wir dürfen also annehmen, daß w von vornherein diesen Verlauf hat. (Diese Annahme ist, wie man leicht sieht, unzulässig, wenn M ein Kreis ist.)

Der Bogen zy läßt sich in so kleine Teilbögen einteilen, daß es zu jedem von ihnen ein ihn im Inneren enthaltendes, x nicht enthaltendes Element gibt. Setzen wir $e=E_1$ und bezeichnen wir die soeben eingeführten Elemente nach der Reihenfolge der in ihnen enthaltenden Teilbögen mit $E_2, \ldots, E_k$, so haben wir folgendes erreicht: w ist in k aneinanderschließende Bögen $w_1, \ldots, w_k$ eingeteilt und diese sind ins Innere von Elementen $E_1, \ldots, E_k$ derart eingeschlossen, daß der Anfangspunkt x von w außer dem Element E_1 keinem weiteren E_i angehört. Eine solche Bedeckung von w mit Elementen möge eine „zulässige Bedeckung von der Ordnung k" heißen.

Unser Satz wird bewiesen sein, wenn wir gezeigt haben, daß es, falls $k > 1$ ist, einen zu w äquivalenten Weg gibt, der eine zulässige Bedeckung von der Ordnung $k-1$ gestattet. Denn dann gelangt man nach $k-1$ Schritten zu einem Weg, der die Behauptung erfüllt.

Die Möglichkeit der Verminderung der Ordnung k einer zulässigen Bedeckung durch Übergang zu einem äquivalenten Weg sei bereits für den Fall $k=2$ bewiesen; dann ergibt sie sich für beliebiges k dadurch, daß man für den Teil $w_1 w_2$ von w und seine aus E_1 und E_2 bestehende zulässige Bedeckung von der Ordnung 2 die Verminderung der Ordnung vornimmt: man ersetzt $w_1 w_2$ durch einen äquivalenten Weg w_1', der eine Bedeckung der Ordnung 1 gestattet, sich also in ein Element E_1' einschließen läßt. Dann ist der Weg w', der aus w bei der Ersetzung von $w_1 w_2$ durch w_1' entsteht, mit w äquivalent, und die Elemente $E_1', E_3, \ldots, E_k$ bilden eine zulässige Bedeckung von w', deren Ordnung $k-1$ ist. Es genügt also, die Möglichkeit der Verminderung der Ordnung für den Fall $k=2$ zu beweisen.

w sei also in zwei Bögen $xq = w_1$, $qy = w_2$ geteilt, w_1 sei im Innern des Elements E_1, w_2 im Innern des Elements E_2, x sei nicht in E_2 enthalten. Da der Bogen w_1 vom Äußeren ins Innere von E_2 läuft, besitzt er Punkte auf dem Rande R von E_2; p sei der im Sinne der Durchlaufung von w_1 letzte derartige Punkt. Der Bogen pq verläuft also in dem Durchschnitt $E_1 \cdot E_2$. Nun sei T eine topologische Abbildung von $E_1 + E_2$ mit den folgenden Eigenschaften: Außerhalb und auf dem Rande von E_2 ist sie die Identität; E_2 wird durch T auf sich abgebildet, und zwar so, daß $T(q) = y$ ist. T kann durch eine eindeutige stetige Deformation hergestellt werden, indem man in E_2 jeden Punkt geradlinig und gleichförmig im Sinne einer in E_2 gültigen euklidischen Geometrie in der Zeit 1 in seinen Bildpunkt laufen läßt und außerhalb und auf dem Rande von E_2 alle Punkte festhält. Da x außerhalb von E_2 und p auf dem Rande von E_2 liegt, werden diese beiden Punkte dabei festgehalten, und der Bogen xp von w ist daher mit seinem Bild $T(xp)$ äquivalent. Ferner ist der Bogen py von w mit dem Bild $T(pq)$ des Bogens pq äquivalent, da beide Wege die Punkte p und y in E_2 miteinander verbinden. Folglich ist auch der Weg $w = xp + py$ mit dem Weg $w' = T(xp) + T(pq) = T(xq) = T(w_1)$ äquivalent. Da aber w_1 im Inneren des Elements E_1 liegt, liegt $w' = T(w_1)$ im Inneren des Elements $E' = T(E_1)$. w' und E' erfüllen also die Behauptung, und der Hilfssatz I ist damit bewiesen.

(**Bemerkung.** Da man innerhalb eines Elementes zwei Punkte immer durch einen doppelpunktfreien Weg verbinden kann, folgt aus dem Hilfssatz, daß man zu jedem Weg w einen äquivalenten, doppelpunktfreien

Weg w' mit denselben Anfangs- und Endpunkten finden kann. Die Konstruktion eines solchen w' ist, wie man leicht sieht, bereits durch eine beliebig kleine Abänderung von w möglich, wenn $n > 2$ ist; dies ist jedoch, wie man an einfachen Beispielen sieht, im Fall $n = 2$ im allgemeinen nicht möglich, vielmehr muß man in diesem Fall Konstruktionen von der Art vornehmen, wie sie im Beweis des Hilfssatzes verwendet wurden.)

2. Ist F eine abgeschlossene Teilmenge eines k-dimensionalen Elements E^k, und ist in F eine Abbildung f auf die n-dimensionale Mannigfaltigkeit μ definiert, so läßt sich diese im allgemeinen — im Gegensatz zu dem Spezialfall, in dem μ der euklidische Raum ist[26]) — nicht auf das ganze Element E^k erweitern. Denn ist z. B. F eine Kreislinie, so ist in der Erweiterbarkeit auf E^k die Erweiterbarkeit auf eine von F begrenzte Kreisscheibe enthalten, und dies bedeutet (siehe § 1), daß der geschlossene Weg $f(F)$ in μ zusammenziehbar ist, was nicht der Fall zu sein braucht. Jedoch ist stets wenigstens eine Erweiterung von f „im Kleinen" möglich; es gilt nämlich

Hilfssatz IIa. Ist in der abgeschlossenen Teilmenge F des k-dimensionalen euklidischen Elements E^k eine Abbildung f auf die Mannigfaltigkeit μ definiert, so läßt sich f auf eine gewisse Umgebung von F erweitern. Dabei wird unter einer Umgebung von F eine F enthaltende, relativ zu E^k offene Teilmenge von E^k verstanden.

Beweis. In μ wird eine bestimmte Metrik mit einer Entfernungsfunktion ϱ zugrunde gelegt. Φ sei eine kompakte Menge in μ. Dann läßt sich jeder positiven Zahl b eine positive (von Φ abhängige) Zahl $\delta(b)$ mit folgender Eigenschaft zuordnen: Jede Punktmenge, die wenigstens einen Punkt von Φ enthält und deren Durchmesser $< \delta(b)$ ist, läßt sich in ein Element mit einem Durchmesser $< b$ einschließen. Gäbe es nämlich zu einem positiven b keine derartige Zahl $\delta(b)$, so gäbe es zu jedem positiven δ_i eine einen Punkt p_i von Φ enthaltende Menge m_i mit einem Durchmesser $< \delta_i$, die sich nicht in ein Element mit einem Durchmesser $< b$ einschließen ließe. Ist aber $\delta_1, \delta_2, \ldots$ eine gegen 0 konvergierende Folge, so hat die zugehörige Punktfolge $p_1, p_2, \ldots$ wegen der Kompaktheit von Φ einen Häufungspunkt p; e sei ein p im Inneren enthaltendes Element mit einem Durchmesser $< b$; ε sei eine so kleine positive Zahl, daß die ganze ε-Umgebung von p im Inneren von e liegt. Ist dann i so groß, daß sowohl $\delta_i < \frac{1}{2}\varepsilon$ als $\varrho(p, p_i) < \frac{1}{2}\varepsilon$ ist, so liegt die ganze Menge m_i im

[26]) Diese Möglichkeit der Erweiterung einer Abbildung folgt unmittelbar aus der Möglichkeit, den Definitionsbereich einer stetigen Funktion zu erweitern. Man vgl. Teil I, Fußnote [17]), und die entsprechende Stelle im Text.

Inneren von e, im Gegensatz zu der Annahme, daß sich m_i in kein Element mit einem Durchmesser $< b$ einschließen lasse.

Eine Funktion $\delta(b)$ existiert also. Aus ihrer Definition ergibt sich, daß immer $\delta(b) \leqq b$ ist. Der Existenzbeweis bleibt gültig für $b = +\infty$; der Sinn der Zahl $\delta^* = \delta(\infty)$ ist der, daß jede Menge, die einen Punkt Φ enthält und deren Durchmesser $< \delta^*$ ist, in ein Element eingeschlossen werden kann.

Wir setzen nun, wenn Φ die Bildmenge $f(F)$ und, wie oben eingeführt, k die Dimension von E^k ist:

$$d_k = \delta^*, \qquad d_\varkappa = \delta\left(\frac{1}{3} d_{\varkappa+1}\right) \qquad (\varkappa = k-1, k-2, \ldots, 1, 0).$$

E^k denken wir uns als Simplex mit einer euklidischen Metrik. Infolge der gleichmäßigen Stetigkeit der Abbildung f läßt sich eine Zahl $c > 0$ so bestimmen, daß für zwei Punkte x_1, x_2 von F, deren Entfernung $< c$ ist, stets $\varrho(f(x_1), f(x_2)) < d_0$ ist.

E^k zerlegen wir in Teilsimplexe, deren Durchmesser $< \frac{1}{3} c$ sind; $E_1^k, E_2^k, \ldots$ seien diejenigen von ihnen, die im Inneren oder auf dem Rande wenigstens je einen Punkt von F enthalten. Die Menge der (relativ zu E^k) inneren Punkte der Vereinigungsmenge $F_k = \sum\limits_i E_i^k$ bildet eine Umgebung von F. Unsere Aufgabe ist daher gelöst, sobald wir $f(F)$ zu einer Abbildung $f(F_k)$ ergänzt haben.

Mit $E_i^\varkappa$ $(\varkappa = 0, 1, \ldots, k-1;\ i = 1, 2, \ldots)$ bezeichnen wir die $\varkappa$-dimensionalen Randsimplexe der E_i^k, mit $F_\varkappa$ die Vereinigungsmenge von F und $\sum\limits_i E_i^\varkappa$ (worin bei festem $\varkappa$ über alle dabei vorkommenden i zu summieren ist). Wir werden f der Reihe nach für $F_0, F_1, \ldots, F_{k-1}, F_k$ erklären.

Ist E_i^0 ein Eckpunkt, der nicht zu F gehört, in dem also f nicht von vornherein definiert ist, so wählen wir willkürlich einen Punkt x von F, der einem E_i^k angehört, an welchem E_i^0 Ecke ist, und setzen: $f(E_i^0) = f(x)$; damit haben wir $f(F_0)$ erklärt. Sind dann y_1, y_2 zwei Punkte von F_0, die einem E_i^k angehören, so gibt es jedenfalls zwei Punkte x_1, x_2 von F (die nicht von den y_1, y_2 verschieden zu sein brauchen), deren Abstand voneinander $< c$ ist, derart, daß $f(y_1) = f(x_1)$, $f(y_2) = f(x_2)$ ist; folglich ist $\varrho(f(y_1), f(y_2)) < d_0$. Aus $y_1 \subset F_0 \cdot E_i^k$, $y_2 \subset F_0 \cdot E_i^k$ (für irgendein i, das aber in beiden Relationen dasselbe ist) folgt also immer $\varrho(f(y_1), f(y_2)) < d_0$.

Es sei nun bereits $f(F_\varkappa)$ so erklärt, daß aus $y_1 \subset F_\varkappa \cdot E_i^k$, $y_2 \subset F_\varkappa \cdot E_i^k$ immer $\varrho(f(y_1), f(y_2)) < d_\varkappa$ folgt. Dann hat, wenn wir ein bestimmtes $E_i^{\varkappa+1}$ betrachten, die Bildmenge $f(F_\varkappa \cdot E_i^{\varkappa+1})$ einen Durchmesser $< d_\varkappa$, sie läßt

sich also infolge der Definition von $d_\varkappa$ und der Definition der δ-Funktion in ein Element e einschließen, dessen Durchmesser $< \frac{1}{3} d_{\varkappa+1}$ ist. Unter Zugrundelegung einer euklidischen Metrik in e läßt sich die Abbildung $f(F_\varkappa \cdot E_i^{\varkappa+1})$ zu einer Abbildung $f(E_i^{\varkappa+1})$ auf eine Teilmenge von e stetig erweitern[26]). Die so in den verschiedenen $E_i^{\varkappa+1}$ erklärten Abbildungen schließen stetig aneinander, da f in allen Simplexen $E_i^\varkappa$ ja schon bei dem vorigen Schritt erklärt worden war; somit ist eine stetige Abbildung $f(F_{\varkappa+1})$ erklärt. Wenn wir noch gezeigt haben, daß dabei aus $y_1 \subset F_{\varkappa+1} \cdot E_i^k$, $y_2 \subset F_{\varkappa+1} \cdot E_i^k$ immer $\varrho(f(y_1), f(y_2)) < d_{\varkappa+1}$ folgt, so ergibt sich durch Induktion die Möglichkeit der Erklärung von $f(F_k)$, womit die Behauptung bewiesen sein wird.

Es ist also noch zu zeigen, daß bei der soeben erklärten Abbildung $f(F_{\varkappa+1})$ aus $y_1 \subset F_{\varkappa+1} \cdot E_i^k$, $y_2 \subset F_{\varkappa+1} \cdot E_i^k$ immer $\varrho(f(y_1), f(y_2)) < d_{\varkappa+1}$ folgt. Ist y ein Punkt von $F_{\varkappa+1} \cdot E_i^k$, der in keinem $E_i^{\varkappa+1}$ liegt, so gehört y zu F, also auch zu $F_\varkappa$; ein beliebiger Eckpunkt z von E_i^k gehört auch zu $F_\varkappa$; mithin ist $y \subset F_\varkappa \cdot E_i^k$, $z \subset F_\varkappa \cdot E_i^k$, folglich $\varrho(f(y), f(z)) < d_\varkappa$, und da stets $\delta(b) \leqq b$ ist, auch $\varrho(f(y), f(z)) < \frac{1}{3} d_{\varkappa+1}$. Ist y ein Punkt von $F_{\varkappa+1} \cdot E_i^k$, der in einem $E_i^{\varkappa+1}$ liegt, so ist, da die Bildmenge $f(E_i^{\varkappa+1})$ in dem Element e enthalten ist, dessen Durchmesser $< \frac{1}{3} d_{\varkappa+1}$ ist, $\varrho(f(y), f(z)) < \frac{1}{3} d_{\varkappa+1}$ für jeden Eckpunkt z von $E_i^{\varkappa+1}$. In jedem Fall gibt es zu dem Punkt y aus $F_{\varkappa+1} \cdot E_i^k$ einen Eckpunkt z von E_i^k mit $\varrho(f(y), f(z)) < \frac{1}{3} d_{\varkappa+1}$. Zu zwei derartigen Punkten y_1, y_2 gibt es demnach zwei derartige Eckpunkte z_1, z_2 von E_i^k. Da z_1 und z_2 zu $F_\varkappa \cdot E_i^k$ gehören, ist $\varrho(f(z_1), f(z_2)) < d_\varkappa \leqq \frac{1}{3} d_{\varkappa+1}$. Da somit jede der drei Entfernungen $\varrho(f(y_1), f(z_1))$, $\varrho(f(z_1), f(z_2))$, $\varrho(f(z_2), f(y_2))$ kleiner als $\frac{1}{3} d_{\varkappa+1}$ ist, ist $\varrho(f(y_1), f(y_2)) < d_{\varkappa+1}$, w. z. b. w.

Damit ist der Hilfssatz IIa bewiesen.

Hilfssatz IIb. E^k, F, μ haben dieselben Bedeutungen wie im Hilfssatz IIa; in F sei aber nicht nur eine Abbildung f, sondern eine für $0 \leqq t \leqq 1$ stetig von dem Parameter t abhängende Schar von Abbildungen f_t auf μ gegeben. Dann läßt sich die ganze Abbildungsschar stetig auf eine Umgebung von F erweitern.

Beweis. Im euklidischen R^{k+1} sei ein rechtwinkliges x^1-x^2-...-x^{k+1}-Koordinatensystem eingeführt. E^k sei durch ein Simplex in dem durch die Gleichung $x^{k+1} = 0$ ausgezeichneten R^k gegeben. Ist $x \subset R^k$, so bezeichne x_t den Punkt, dessen senkrechte Projektion auf R^k der Punkt x und dessen x^{k+1}-Koordinate die Zahl t ist. $\bar{F}$ sei die Menge der Punkte x_t im R^{k+1}, für die $x \subset F$, $0 \leqq t \leqq 1$ ist, $\bar{E}^{k+1}$ der Quader im R^{k+1}, der von den Punkten x_t mit $x \subset E^k$, $0 \leqq t \leqq 1$ gebildet wird. Durch $f(x_t) = f_t(x)$ für $x_t \subset \bar{F}$ ist eine Abbildung $f(\bar{F})$ definiert; sie läßt sich nach Hilfssatz IIa zu einer Abbildung einer Umgebung $\bar{U}$ von $\bar{F}$ in $\bar{E}^{k+1}$ ergänzen. Es gibt

eine Umgebung U von F in E^k derart, daß aus $x \subset U$, $0 \leq t \leq 1$ folgt: $x_t \subset \bar{U}$. In U wird durch $f_t(x) = f(x_t)$ die gewünschte Erweiterung von $f_t(F)$ geleistet.

Hilfssatz IIc. F sei eine abgeschlossene Teilmenge von E^k. In E^k sei eine Abbildung f_0 auf μ, in F eine an f_0 anschließende stetige Abbildungsschar f_t für $0 \leq t \leq 1$ definiert, die in den zum Rande R von E^k gehörigen Punkten von F für alle t-Werte mit f_0 übereinstimmt. Dann läßt sich in E^k eine an f_0 anschließende stetige Abbildungsschar für $0 \leq t \leq 1$ erklären, die in F mit der Schar f_t und auf R für alle t-Werte mit f_0 übereinstimmt.

Beweis. Es sei $f_t(R) = f_0(R)$, $F_1 = F + R$. Nach IIb läßt sich die Schar $f_t(F_1)$ auf eine Umgebung U von F_1 erweitern. ε sei so klein, daß die ε-Umgebung von F_1 in U liegt. Ist y ein Punkt von E^k, so bezeichne $r(y)$ die Entfernung des Punktes y von der Menge F_1. Die folgendermaßen erklärte Abbildungsschar $g_t(E^k)$ hat die gewünschten Eigenschaften:

$$\begin{aligned} &\text{Ist } r(y) \geqq \varepsilon, \text{ so ist } g_t(y) = f_0(y) && \text{für } 0 \leqq t \leqq 1; \\ &\text{ist } r(y) \leqq \varepsilon, \text{ so ist } g_t(y) = f_t(y) && \text{für } 0 \leqq t \leqq 1 - \tfrac{r(y)}{\varepsilon} \\ &\qquad\text{und } g_t(y) = f_{1-\frac{r(y)}{\varepsilon}}(y) && \text{für } 1 - \tfrac{r(y)}{\varepsilon} \leqq t \leqq 1. \end{aligned}$$

3. In der euklidischen Ebene ist jede Kreislinie zusammenziehbar; entfernt man aber den Mittelpunkt eines Kreises aus der Ebene, so entsteht eine einem Zylinder homöomorphe Fläche, auf der derselbe Kreis nicht mehr zusammenziehbar ist. Jedoch kann die Zusammenziehbarkeit eines geschlossenen Weges in einer Mannigfaltigkeit durch Herausnehmen eines Punktes aus dieser nur dann zerstört werden, wenn die Dimensionszahl der Mannigfaltigkeit 2 ist; es gilt nämlich:

Hilfssatz III. Ist $n > 2$, w ein in der n-dimensionalen Mannigfaltigkeit μ zusammenziehbarer Weg, ξ ein nicht auf w liegender Punkt von μ, μ' die durch Herausnahme von ξ aus μ entstehende Mannigfaltigkeit, so ist w auch in μ' zusammenziehbar.

Beweis. Die Zusammenziehbarkeit von w in μ bedeutet, daß w das Bild des Randes R bei einer Abbildung g einer Kreisscheibe, oder, was dasselbe bedeutet, einer Dreieckscheibe T auf μ ist. Falls das Bild $g(T)$ den Punkt ξ nicht enthält, ist $g(T) \subset \mu'$, w also in M' zusammenziehbar; falls $\xi \subset g(T)$ ist, besteht unsere Aufgabe darin, eine auf R mit g übereinstimmende Abbildung h von T auf μ zu konstruieren, so daß ξ nicht in der Bildmenge $h(T)$ enthalten ist.

Es sei also $\xi \subset g(T)$; X sei die Originalmenge von ξ in T. e sei ein ξ im Inneren enthaltendes Element; ε sei eine so kleine positive Zahl, daß

die ε-Umgebung von ξ in e enthalten ist. Infolge der gleichmäßigen Stetigkeit der Abbildung $g(T)$ gibt es eine positive Zahl δ von der Eigenschaft, daß das Bild jeder Punktmenge von T, deren Durchmesser $< \delta$ ist, einen Durchmesser hat, der $< \varepsilon$ ist. Wir nehmen mit T eine Unterteilung in Dreiecke $t_1, t_2, \ldots$ vor, deren Durchmesser $< \delta$ und kleiner als der Abstand zwischen R und X sind; dann liegt das Bild jedes einen Punkt von X enthaltenden Dreiecks t_i im Inneren von e, und die Vereinigungsmenge Q derjenigen t_i, die wenigstens je einen Punkt von X im Inneren oder auf dem Rande enthalten, ist fremd zu dem Rande R. Wir tilgen nun die Abbildung g im Inneren von Q; auf R bleibt sie also bestehen. Außerhalb und auf dem Rande von Q setzen wir $h = g$. Jedem im Inneren von Q liegenden Eckpunkt eines t_i ordnen wir als Bild bei der Abbildung h einen beliebigen, von ξ verschiedenen, inneren Punkt von e zu. Darauf erweitern wir auf jeder Seite eines t_i, deren innere Punkte im Inneren von Q liegen, die in ihren Ecken schon erklärte Abbildung zu einer solchen Abbildung h der ganzen Seite in das Innere von e, daß ξ nicht zu der dabei entstehenden Bildmenge gehört; dies ist möglich, da $n > 1$ ist[27]). Schließlich erweitern wir in jedem zu Q gehörigen t_i die auf seinen Seiten schon erklärte Abbildung zu einer solchen Abbildung h des ganzen t_i in das Innere von e, daß ξ nicht in der Bildmenge $h(t_i)$ enthalten ist; dies ist möglich, da $n > 2$ ist[27]). Damit ist $h(T)$ in der gewünschten Weise konstruiert.

Bemerkung. Man sieht übrigens leicht, daß der Hilfssatz III trivialerweise auch für $n = 1$ gilt, da in diesem Falle jeder zusammenziehbare geschlossene Weg in sich selbst zusammenziehbar ist, so daß die Zusammenziehbarkeit durch Herausnahme eines nicht auf ihm gelegenen Punktes aus M nicht gestört wird.

Durch mehrmalige Anwendung des Hilfssatzes III wird dieser erweitert zu:

Hilfssatz IIIa. Die Aussage des Hilfssatzes III behält ihre Gültigkeit, wenn man an Stelle des Punktes ξ eine Menge von endlich vielen Punkten $\xi_1, \xi_2, \ldots, \xi_m$ und an Stelle von μ' die durch Herausnahme dieser Punkte aus μ entstandene Mannigfaltigkeit betrachtet.

§ 6.

Die Bestimmung der im § 4 definierten Mindestzahlen für $n \neq 2$.

Wir beginnen mit dem Beweis des am Schluß des § 4 angekündigten Satzes, daß die in der Klasse $\mathfrak{K}_\xi$ erreichbare Mindestzahl von Original-

[27]) Teil I, § 2, Hilfssatz I.

punkten des Punktes ξ gleich der wesentlichen Schichtenzahl in ξ ist, falls die Dimensionenzahl nicht 2 ist, für einen Spezialfall:

Satz XIIIa. *Es sei $n \neq 2$. Das n-dimensionale Element E sei auf die n-dimensionale Mannigfaltigkeit μ so abgebildet, daß die Originalmenge des Punktes ξ von μ aus zwei im Inneren von E gelegenen Punkten besteht, deren Verbindungswege in E zusammenziehbare Bilder in μ haben. Dann läßt sich die Abbildung durch eine stetige Abänderung, die nur die Bilder innerer Punkte von E verrückt, in eine Abbildung überführen, bei der ξ einen einzigen Originalpunkt hat.*

Beweis. Ist $n = 1$, so wird durch die nach Voraussetzung mögliche Zusammenziehung des Bildes der Verbindungsstrecke der Originalpunkte x_1, x_2 von ξ auf den Punkt ξ die Abbildung $f(E)$ ohne Änderung in den beiden Randpunkten von E stetig in eine Abbildung $f_1(E)$ übergeführt, bei der die Strecke $x_1 x_2$ die Originalmenge von ξ ist. f_1 läßt sich (siehe Satz XII) ohne Änderung in den Endpunkten von E weiter stetig in eine Abbildung überführen, bei der ξ einen einzigen Originalpunkt hat.

Es sei $n > 2$. e_1, e_2 seien ξ enthaltende Elemente, und e_2 liege im Inneren von e_1. E denken wir uns als Simplex, v sei die Verbindungsstrecke der Originalpunkte x_1, x_2 von ξ. Wir werden zuerst zeigen, daß man die Abbildung f unter Festhaltung des Randbildes so abändern kann, daß $x_1 + x_2$ die Originalmenge von ξ bleibt und daß das Bild von v ganz im Inneren von e_1 liegt.

Das letztere sei also noch nicht der Fall. Dann gibt es bei Durchlaufung von v in der Richtung $x_1 x_2$ einen ersten Punkt p und einen letzten Punkt q mit Bildern auf dem Rande r von e_2; wir bezeichnen die Strecken $x_1 p$, pq, $q x_2$ der Reihe nach mit v_1, v_2, v_3; ferner sei w ein Weg von $f(p)$ nach $f(q)$ auf r (w existiert, da $n > 2$ ist). Der geschlossene Weg $f(v_1)^{-1} f(v_3)^{-1} w^{-1}$ ist zusammenziehbar, weil er in dem Element e_1 liegt, der geschlossene Weg $f(v_2) f(v_3) f(v_1)$ ist zusammenziehbar, weil der Weg $f(v_1) f(v_2) f(v_3) = f(v)$ nach Voraussetzung zusammenziehbar ist; folglich ist auch $f(v_2) w^{-1} = f(v_2) f(v_3) f(v_1) f(v_1)^{-1} f(v_3)^{-1} w^{-1}$ zusammenziehbar. Da ξ auf v_2 keinen Originalpunkt hat und da w auf r verläuft, liegt ξ nicht auf dem Wege $f(v_2) w^{-1}$; dieser ist daher nach Hilfssatz III nicht nur in μ, sondern auch in der durch Herausnahme von ξ aus μ entstehenden Mannigfaltigkeit μ' zusammenziehbar. w ist also mit $f(v_2)$ in μ' äquivalent (siehe § 5, Anfang), und man kann $f(v_2)$ unter Festhaltung seiner Endpunkte $f(p), f(q)$ innerhalb μ' in w deformieren. Da die Strecke v_2 keinen Doppelpunkt hat, entspricht dieser Deformation eine stetige Schar eindeutiger Abbildungen $f_t(v_2)$ von v_2 in die Mannigfaltigkeit μ' mit $0 \leq t \leq 1$, $f_0(v_2) = f(v_2)$, $f_1(v_2) = w$.

Es sei nun E' ein im Inneren von E gelegenes Element mit folgenden Eigenschaften: p und q liegen auf dem Rande, die übrigen Punkte von v_2 liegen im Inneren, die von p und q verschiedenen Punkte von v_1 und v_3 liegen im Äußeren von E'. Auf E', $v_2 = F$ und μ' wenden wir den Hilfssatz IIc an; dann gelangen wir zu einer für $0 \leqq t \leqq 1$ stetigen Abbildungsschar $f_t(E')$ mit $f_0 = f$, mit $f_t = f$ am Rande von E' und mit $f_1(v_2) = w$. Fassen wir die f_t als Abbildungen von E' auf die Mannigfaltigkeit μ auf und setzen wir dann $f_t = f$ in $E - E'$ für $0 \leqq t \leqq 1$, so haben wir eine stetige Abbildungsschar $f_t(E)$ mit $f_0 = f$, mit $f_t = f$ am Rande von E, mit $f_1(v) = f(v_1)\,w\,f(v_3)$ und mit der Eigenschaft, daß $x_1 + x_2$ für alle t-Werte die Originalmenge von ξ ist. Da $f_1(v) \subset e_1$ ist, erhält man, während t von 0 bis 1 läuft, eine solche Abänderung von f, wie sie oben (2. Absatz des Beweises) als vorläufiges Ziel hingestellt wurde.

Damit sind wir aber auch im wesentlichen am Ende des Beweises; denn ist ε eine so kleine positive Zahl, daß das durch f_1 gelieferte Bild der ε-Umgebung von v in e_1 liegt, so kann man, da die abgeschlossene Hülle der ε-Umgebung der Strecke v ein Element E'' ist, in E'' f_1 weiter so abändern, daß am Rande von E'' nichts geändert wird und daß ξ schließlich nur noch einen einzigen Originalpunkt hat[22]).

Damit ist der Vorbereitungssatz XIIIa bewiesen, und wir haben jetzt die Gleichung $\sigma_\xi = s_\xi$, die ja das Ziel unserer gegenwärtigen Überlegungen ist, für $n \neq 2$ in voller Allgemeinheit zu beweisen. Hierzu haben wir, wenn die Abbildung f im Punkt ξ kompakt und wenn dort die wesentliche Schichtenzahl s_ξ ist, eine zu der „Klasse $\mathfrak{K}_\xi$ in bezug auf ξ", die durch f bestimmt ist, gehörige Abbildung f_1 anzugeben, bei der ξ genau s_ξ Originalpunkte besitzt. Wir werden etwas mehr beweisen; wir werden nämlich den Bereich der zulässigen Abänderungen von f dadurch noch einschränken, daß wir die „Klasse in bezug auf ξ" [20]) durch die „Klasse $\mathfrak{K}$" [20]), zu der f gehört, ersetzen. In den Beweisen der folgenden Sätze dieses Paragraphen wird die Zugehörigkeit zu $\mathfrak{K}$ stets dadurch sichergestellt sein, daß alle vorkommenden Abänderungen von f sich immer nur auf kompakte Teile von M beziehen und außerhalb dieser Teile nichts ändern.

Satz XIIIb. *Es sei $n \neq 2$, M und μ seien n-dimensionale Mannigfaltigkeiten, f sei eine Abbildung von M auf μ, die im Punkte ξ kompakt ist; dann gibt es in der durch f bestimmten Klasse $\mathfrak{K}$ eine Abbildung f_1, bei der die Originalmenge X von ξ keine unwesentliche Schicht enthält und bei der jede wesentliche Schicht von X aus nur einem Punkt besteht.*

Beweis. Es sei zunächst $n > 2$. Da sich jede Abbildung durch eine kleine Abänderung „fastglatt" in einem gegebenen Punkt machen läßt[22]), dürfen wir annehmen, daß die Originalmenge X von ξ nur aus endlich

vielen Punkten besteht. x_1, x_2 seien Punkte einer Schicht von X; dann gibt es einen Weg u von x_1 nach x_2, dessen Bild $f(u)$ zusammenziehbar ist; nach Hilfssatz I gibt es ein x_1 und x_2 im Inneren enthaltendes Element, so daß jeder Weg v, der x_1 mit x_2 in dem Element verbindet, äquivalent mit u, daß daher sein Bild auch zusammenziehbar ist. Falls dieses Element außer x_1 und x_2 noch andere Punkte von X enthält, ersetzen wir es durch ein x_1 und x_2 enthaltendes Teilelement E, welches die anderen Punkte von X in seinem Äußeren läßt. Indem wir auf E den Satz XIIIa anwenden, ersetzen wir x_1 und x_2 durch einen einzigen Originalpunkt von ξ; wir vermindern also die Anzahl der Originalpunkte. Dies wiederholen wir so lange, bis jede Schicht nur noch einen einzigen Punkt enthält.

Wir haben nun noch diejenigen von diesen Punkten, die unwesentliche Schichten darstellen, zu beseitigen. Ist x ein solcher unwesentlicher Punkt, und ist f orientierbar, so hat die Abbildung eines kleinen, x enthaltenden Elementes im Punkt ξ den Grad 0, und man kann die Abbildung in diesem Element, ohne sie auf seinem Rande zu ändern, in eine Abbildung überführen, bei der ξ nicht mehr zu der Bildmenge gehört[28]). So werden, wenn f orientierbar ist, alle unwesentlichen Schichten beseitigt. Ist f nicht orientierbar, so hat die Abbildung eines kleinen, x enthaltenden Elementes e im Punkte ξ einen geraden Grad $2c$. Wir ersetzen f im Inneren von e mittels stetiger Abänderung, die am Rande von e wieder nichts ändert, durch eine solche Abbildung, daß ξ zwei Originalpunkte x_1, x_2 erhält und daß die Abbildungen von Umgebungen U_1 bzw. U_2 dieser beiden Punkte in ξ je den Grad c haben, wobei die Orientierungen von U_1 und U_2 durch eine Orientierung von e induziert sind[29]). Nun sei i ein geschlossener, die Orientierung umkehrender Weg durch x_2, dessen Bild zusammenziehbar ist; ist u ein x_1 mit x_2 innerhalb e verbindender Weg, so liefert die Fortsetzung einer bestimmten Orientierung von U_1 längs ui die entgegengesetzte Orientierung in U_2 wie die Fortsetzung längs u. Da das Bild von u in der Umgebung von ξ verläuft, ist es zusammenziehbar, und mithin ist auch das Bild von ui zusammenziehbar. Wir schließen jetzt, was nach Hilfssatz I möglich ist, x_1 und x_2 in ein solches Element E ein, daß jeder x_1

[28]) Teil I, § 1, Satz IXa.

[29]) Deutet man die in einer Umgebung von ξ mit ξ als Nullpunkt eingeführten euklidischen Koordinaten der Bildpunkte als die Komponenten von Vektoren, die den Punkten von e zugeordnet sind, so ist die hier zu lösende Aufgabe identisch mit der Aufgabe, ein in e gegebenes Vektorfeld, das am Rande den Index $2c$ hat, durch eine Abänderung im Inneren von e durch ein anderes Vektorfeld, das dort genau zwei Nullstellen mit den Indexen c hat, zu ersetzen; diese „Randwertaufgabe" ist lösbar; siehe H. Hopf, Abbildungsklassen n-dimensionaler Mannigfaltigkeiten, § 5, Nr. 4, Math. Annalen **96** (1926).

mit x_2 in E verbindende Weg v äquivalent mit ui ist; dann ist auch das Bild von v zusammenziehbar. Da die Fortsetzung der Orientierung von U_1 nach U_2 längs v dasselbe Ergebnis hat wie die Fortsetzung längs dem Weg ui, also das entgegengesetzte Ergebnis wie die Fortsetzung längs u, hat die Abbildung einer der Umgebungen U_1, U_2, wenn man sie als Teile des Elementes E orientiert, in ξ den Grad $+c$, die der anderen den Grad $-c$; die Abbildung von E hat also in ξ den Grad 0. Da wir wieder annehmen dürfen, daß E außer x_1 und x_2 keinen Originalpunkt von ξ enthält, können wir nach Satz XIIIa durch eine Änderung im Inneren von E die Abbildung in eine solche überführen, die in E nur einen Originalpunkt x_0 von ξ besitzt. Der Grad in ξ ändert sich bei der Abänderung nicht; mithin hat die Abbildung einer Umgebung U_0 von x_0 den Grad 0, und man kann daher den Originalpunkt x_0 von ξ mittels einer weiteren Abänderung der Abbildung in U_0 beseitigen[28]). — Damit ist unser Satz für $n > 2$ bewiesen.

Ist $n = 1$, so unterscheiden wir zwei Fälle, je nachdem M ein Kreis oder eine Gerade ist. Im ersten Fall läßt sich, falls μ eine Gerade ist, die Bildmenge auf einen beliebigen Punkt von μ zusammenziehen, die Originalmenge eines gegebenen Punktes ξ läßt sich also überhaupt beseitigen; falls M und μ Kreise sind, läßt sich die Abbildung f, wenn ihr Grad $c \neq 0$ ist, in eine monotone c-malige Umlaufung von μ deformieren; bei einer solchen besteht jede der c wesentlichen Schichten für jeden Punkt ξ aus genau einem Punkt; ist $c = 0$, so läßt sich die Bildmenge wieder auf einen Punkt zusammenziehen, und die Originalmengen aller anderen Punkte sind leer.

Ist M eine Gerade, so ändern wir wie im Fall $n > 2$ f zunächst so ab, daß ξ nur endlich viele Originalpunkte hat. Sind x_1, x_2 Punkte einer Schicht, so ist, da auf der Geraden je zwei Verbindungswege zwischen zwei Punkten einander äquivalent sind, das Bild der Strecke $x_1 x_2$ zusammenziehbar. Wir ändern f stetig ab, indem wir es auf den Punkt ξ zusammenziehen. Haben wir dies getan, so können wir wieder wie im Beweis des vorigen Satzes die Abbildung im Inneren eines x_1 und x_2 im Inneren enthaltenden Intervalles, ohne sie in dessen Endpunkten zu ändern, stetig in eine Abbildung überführen, bei der ξ in diesem Intervall nur noch einen Originalpunkt enthält (siehe Satz XII). Auf diese Weise erreichen wir, wie im Fall $n > 2$, daß jede Schicht der Originalmenge von ξ nur einen Punkt enthält, und dieser läßt sich, wie bei einer orientierbaren Abbildung im Fall $n > 2$, beseitigen, falls die durch ihn repräsentierte Schicht unwesentlich ist.

Damit ist der Satz XIIIb bewiesen; er läßt sich noch verschärfen zu

Satz XIIIc. *Es sei $n \neq 2$, M und μ seien n-dimensionale Mannigfaltigkeiten, die Abbildung f von M auf μ sei in den Punkten $\xi^1, \xi^2, \ldots, \xi^m$ kompakt; dann läßt sich f innerhalb der Klasse $\mathfrak{K}$ stetig in eine Abbil-*

dung f_1 *überführen, bei der keine der Originalmengen* X^k *der Punkte* ξ^k $(k = 1, 2, \ldots, m)$ *eine unwesentliche Schicht enthält und bei der jede wesentliche Schicht einer Menge* X^k *nur aus einem einzigen Punkt besteht.*

Beweis. Es sei $n > 2$. Wir dürfen annehmen, daß jede der Mengen X^k von vornherein nur aus endlich vielen Punkten besteht[33]). Entfernt man einige der Punkte ξ^k aus μ und die entsprechenden Mengen X^k aus M, so bleiben Mannigfaltigkeiten μ' und M' übrig, wobei letztere auf erstere so abgebildet ist, daß diese Abbildung in den übriggebliebenen Punkten ξ^k kompakt ist; für deren Originalmengen hat sich überdies auf Grund des Hilfssatzes III die Schichtenzerlegung nicht geändert.

M_1, μ_1 seien die Mannigfaltigkeiten, die durch Herausnahme aller ξ^k und X^k mit Ausnahme von ξ^1 und X^1 entstehen. Wenden wir dann den Satz XIII b zunächst auf M_1, μ_1 und ξ^1 an, so wird die Behauptung unseres Satzes erfüllt, ohne daß an den Originalmengen der Punkte $\xi^2, \xi^3, \ldots, \xi^m$ etwas geändert würde. Darauf wenden wir den Satz XIII b auf M_2, μ_2 und ξ^2 an, wobei M_2 und μ_2 die Mannigfaltigkeiten sind, die durch Entfernung der (jetzt vorliegenden) Originalmengen von $\xi^1, \xi^3, \ldots, \xi^m$ aus M und der Punkte $\xi^1, \xi^3, \ldots, \xi^m$ aus μ entstehen; dabei wird an keiner der Mengen $X^1, X^3, \ldots, X^m$ etwas geändert, insbesondere wird das für ξ^1 bereits erzielte Ergebnis nicht wieder zerstört, und die Behauptung unseres Satzes wird also für ξ^1 und ξ^2 erfüllt. So fortfahrend erfüllen wir sie schließlich für alle ξ^k.

Ist $n = 1$ und M ein Kreis, so erfüllt, wie im Beweis des vorigen Satzes, diejenige durch stetige Abänderung von f entstandene Abbildung f_1, die M auf einen einzigen Punkt abbildet, oder die eine monotone mehrmalige Durchlaufung des Bildkreises μ darstellt, die Behauptung. Ist M eine Gerade, so führt die im Beweis des vorigen Satzes für den Punkt ξ angegebene Konstruktion zum Ziel, wenn man sie nacheinander für $\xi^1, \xi^2, \ldots$ durchführt; denn wendet man sie auf ξ^k an, so besteht ihr erster Schritt in der Zusammenziehung des Bildes einer Strecke auf den Punkt ξ^k, wobei die Originalmenge keines von ξ^k verschiedenen Punktes vermehrt wird, und ihr zweiter (und letzter) Schritt — das Ersetzen dieser Strecke durch einen einzigen Originalpunkt von ξ^k — ändert die Abbildung nur in einer beliebig kleinen Umgebung von ξ^k. Unsere Behauptung gilt also auch für $n = 1$.

Wir weisen noch auf eine Folgerung aus den Sätzen XIII c und VII a hin:

Satz XIII d. *Ist* $n \neq 2$, f *eine überall kompakte Abbildung von* M *auf* μ *mit einem von* 0 *verschiedenen Absolutgrad und mit dem Index* j, *so gibt es in der Klasse von* f *eine Abbildung, bei der jeder von endlich*

vielen, in μ willkürlich gegebenen Punkten $\xi^1, \xi^2, \ldots, \xi^m$ genau j Originalpunkte hat.

Wenn im Satz XIIIb der Absolutgrad in ξ $a_\xi = 0$, wenn also auch die wesentliche Schichtenzahl $s_\xi = 0$ ist, so besagt der Satz, daß es eine Abbildung f_1 in $\mathfrak{K}$ gibt, bei der ξ keinen Originalpunkt hat, also nicht zu der Bildmenge gehört. Diese Tatsache verallgemeinern wir, indem wir damit die zweite am Schluß des § 4 ausgesprochene Behauptung, nämlich: „${}_n a_\xi = a_\xi$ für $n \neq 2$", beweisen, zu folgendem Satz:

Satz XIV. *Ist $n \neq 2$, ist die Abbildung f von M auf μ in den Punkten ξ^k kompakt und hat sie dort die Absolutgrade a^k $(k = 1, 2, \ldots, m)$, so gibt es in der durch f bestimmten Klasse $\mathfrak{K}$ eine Abbildung f^*, die in den Punkten ξ^k glatt ist und bei der je eine Umgebung dieser Punkte a^1-, a^2-, ..., a^m-mal glatt bedeckt wird.*

Beweis. Auf Grund von Satz XIIIc dürfen wir annehmen, daß es in jedem der Punkte ξ^k nur wesentliche Schichten gibt und daß jede von diesen aus einem einzigen Punkt besteht. Die Originalmenge X^k von ξ^k bestehe also aus den Punkten $x_1^k, x_2^k, \ldots$, deren jeder eine Schicht darstellt, und a_i^k sei der Beitrag der durch x_i^k dargestellten Schicht. e_i^k sei ein so kleines, x_i^k enthaltendes Element, daß das Bild $f(e_i^k)$ in einer euklidischen Umgebung V^k von ξ^k liegt.

Ist f orientierbar, so ist a_i^k der Betrag des Grades in ξ^k bei der Abbildung $f(e_i^k)$. Man kann daher durch Änderung im Inneren von e_i^k und V^k zu einer Abbildung f' übergehen, die sich am Rande von e_i^k stetig an f anschließt und bei der das Bild $f'(e_i^k)$ eine Umgebung von ξ^k a_i^k-mal glatt bedeckt[29]). Führt man dies für jeden Punkt x_i^k einzeln durch, so gelangt man, da $\sum_i a_i^k = a^k$ ist, zu einer Abbildung f^*, die die Behauptung erfüllt.

Ist f nicht orientierbar, so ist $a_i^k = 1$ und der Grad in ξ^k bei der Abbildung $f(e_i^k)$ ist ungerade. Wir dürfen aber sogar annehmen, daß er 1 ist. Denn ist er zunächst $2c + 1$, so können wir durch eine Abänderung in e_i^k erreichen, daß ξ^k dort zwei Originalpunkte y, z hat und daß die Abbildung einer Umgebung von y in ξ^k den Grad 1, die Abbildung einer Umgebung von z in ξ^k den Grad $2c$ hat[29a]). Dann können wir nach dem Verfahren, mit dem wir im Beweis des Satzes XIIIb die unwesentlichen Schichten im Fall einer nicht orientierbaren Abbildung beseitigt haben, den Punkt z aus der Originalmenge von ξ^k entfernen, so daß man zu den Sätzen XIIIb und XIIIc den Zusatz machen kann: Im Fall einer nicht orientierbaren Abbildung kann die Abbildung f_1 so gewählt werden, daß die Abbildungen der Umgebungen der die einzelnen Schichten repräsentie-

---

[29a]) Dies geschicht analog wie in Fußnote [29]) angegeben.

renden Punkte in ξ bzw. in allen Punkten $\xi^1, \xi^2, \ldots, \xi^m$ sämtlich den Grad 1 haben.

$f(e_i^k)$ habe also in ξ^k den Grad 1. Dann kann man durch Änderung im Inneren von e_i^k und V^k zu einer Abbildung f' übergehen, die sich am Rande von e_i^k stetig an f anschließt und für die die Teilabbildung $f(e_i^k)$ in der Umgebung von ξ^k eineindeutig ist[25]). Durchführung dieses Verfahrens für jeden einzelnen Punkt x_i^k liefert schließlich eine Abbildung f^*, die die Behauptung erfüllt.

Es liegt nun die Frage nahe, wie weit sich die a^k-maligen glatten Bedeckungen, die wir in kleinen Umgebungen der Punkte ξ^k erzielt haben, ausdehnen lassen. Der folgende Satz gibt eine Antwort.

Satz XIVa. *Ist $n \neq 2$, sind G^k $(k = 1, 2, \ldots)$ Gebiete von μ, in deren jedem f kompakt ist, und K^k Teilmengen der G^k, die in diesen kompakt sind, so gibt es in der Klasse von f eine Abbildung f^*, bei der in jeder Menge K^k eine dort überall dichte, offene Teilmenge genau a^k-mal glatt von der Bildmenge bedeckt wird, wobei a^k der Absolutgrad von f in G^k ist.*

Beweis. Ist T ein n-dimensionales Simplex, t ein n-dimensionales Teilsimplex von T, so verstehen wir unter dem „Aufblasen von t auf T" eine eindeutige Deformation von T in sich, bei der alle Randpunkte von T festbleiben, alle Punkte des Zwischengebietes $T - t$ und des Randes von t auf den Rand von T wandern und bei deren Endergebnis t eineindeutig auf T abgebildet ist. Eine solche Deformation läßt sich, wenn T und t gegeben sind, stets in elementarer Weise angeben.

E sei ein Teilelement von G^k, ξ ein innerer Punkt von E, und f sei bereits in seiner Klasse so abgeändert, daß eine in E gelegene Umgebung von ξ a^k-mal glatt bedeckt wird. e sei ein Teilelement dieser Umgebung, das bei einer topologischen Abbilduug von E auf ein Simplex T einem Teilsimplex t von T entspricht. Dann entspricht dem Aufblasen von t auf T eine Deformation D_τ $(0 \leqq \tau \leqq 1)$ von μ, die wir das „Aufblasen von e auf E" nennen; da f in E kompakt und D_τ außerhalb E die Identität ist, sind die Abbildungen $f_\tau = D_\tau f$ überall dort kompakt, wo f kompakt ist, gehören also zu der Klasse von f. Bei dem Endergebnis f_1 wird das ganze Innere von E genau a^k-mal glatt bedeckt.

Sei jetzt E' ein zweites Teilelement von G^k, das mit E innere Punkte gemeinsam habe; die Menge dieser Punkte wird bei f_1 a^k-mal glatt bedeckt. In ihr läßt sich durch Abbildung von E' auf ein Simplex T ein Element e' finden, das man auf E' aufblasen kann; tun wir dies, so wird bei dem Endergebnis f_2 der durch diese Deformation bewirkten Abänderung von f_1, die wieder innerhalb der Klasse von f vor sich geht, das

ganze Innere von E', sowie der nicht auf dem Rande von E' gelegene Teil des Inneren von E a^k-mal glatt bedeckt. Da der Rand von E' in E nirgends dicht ist, wird somit eine offene, in $E + E'$ überall dichte Menge a^k-mal glatt bedeckt.

Ist nun E'' ein drittes Element von G^k, das mit $E + E'$ innere Punkte gemeinsam hat, so läßt sich die a^k-malige glatte Bedeckung ebenso auf eine offene, in $E + E' + E''$ überall dichte Menge erweitern, nämlich auf die Menge derjenigen inneren Punkte von $E + E' + E''$, die nicht auf dem Rande eines dieser Elemente liegen. Dasselbe Ergebnis läßt sich durch wiederholtes Aufblasen geeigneter kleiner Elemente für eine beliebige endliche Anzahl von Elementen $E, E', \ldots, E^{(m)}$ erzielen, von denen jedes mit wenigstens einem vorhergehenden innere Punkte gemeinsam hat. Hieraus folgt, da sich jede Komponente jeder der gegebenen Mengen K^k ins Innere eines derartigen Systems von Elementen einschließen läßt, die Richtigkeit unserer Behauptung.

Betrachten wir noch den Spezialfall, daß M geschlossen ist: in ihm ist f überall in μ kompakt, μ kann also die Rolle eines G^k aus dem eben bewiesenen Satz übernehmen. Ist auch μ geschlossen, so kann μ überdies die Rolle des zugehörigen K^k übernehmen. Ist μ offen, so wird nur ein echter Teil von μ durch die Bildmenge bedeckt und es ist $a = 0$; man kann die Bildmenge mit endlich vielen Elementen $E, E', \ldots, E^{(m)}$ so bedecken, daß E auch Punkte enthält, die nicht zu der Bildmenge gehören, und daß jedes weitere der Elemente mit wenigstens einem vorhergehenden innere Punkte gemeinsam hat. Dann liefert der beim Beweise des vorigen Satzes benutzte Prozeß des sukzessiven Aufblasens, wenn man mit einem 0-mal bedeckten Teilelement e von E beginnt, eine Abbildung f^*, bei der die Bildmenge ganz auf den Rändern der $E, E', \ldots, E^{(m)}$ liegt. Es gilt also:

Satz XIVb. *Ist $n \neq 2$, M geschlossen und f eine Abbildung von M auf μ vom Absolutgrad a, so gibt es in der Klasse von f eine Abbildung f^*, bei der eine offene, in μ überall dichte Menge genau a-mal glatt bedeckt wird. Ist insbesondere $a = 0$, so ist bei f^* die Bildmenge nirgends dicht in μ.*

Die vorstehenden Sätze sind unter der Voraussetzung $n \neq 2$ bewiesen. Jedoch gilt, wie H. Kneser gezeigt hat[3]) [6]), der dem Satz XIVb entsprechende Satz auch, wenn M und μ geschlossene *Flächen* sind. Ferner macht der soeben durchgeführte, auf offene μ bezügliche Teil des Beweises von XIVb keinen Gebrauch von der Voraussetzung $n \neq 2$ (da ja in ihm die a-malige, d. h. 0-malige, glatte Bedeckung eines Teiles von E von selbst vorhanden ist und nicht erst auf Grund früherer Sätze hergestellt zu werden braucht); die Annahme der Geschlossenheit von μ ist für die Gültigkeit des Kneserschen Satzes also unnötig. Mithin gilt

Satz XIVc. *Die Voraussetzung $n \neq 2$ ist im Satz* XIVb *unnötig.*

§ 7.

Die Sonderstellung der Flächenabbildungen.

Wir werden jetzt die am Schluß des § 4 ausgesprochene Behauptung beweisen, daß die Gleichungen $\sigma_\xi = s_\xi$, $\alpha_\xi = a_\xi$, deren Richtigkeit für $n \neq 2$ wir soeben erkannt haben, nicht bei allen Flächenabbildungen gelten. Wir beginnen mit einer Hilfsbetrachtung:

P sei eine euklidische x-y-Ebene; die Punkte der x-Achse bezeichnen wir kurz durch ihre x-Koordinate. U, V seien die Kreise mit dem Radius 1 um die Punkte -1, $+1$. P sei orientiert, und wir verstehen unter U, V die Kreise in ihren positiven, unter U^{-1}, V^{-1} die Kreise in ihren negativen Durchlaufungsrichtungen. E sei eine in einer zweiten orientierten Ebene M gelegene Kreisscheibe, C der Randkreis von E in positiver Durchlaufungsrichtung. Wir teilen C in vier in positiver Richtung aufeinander folgende Bögen b_1, b_2, b_3, b_4 und definieren folgende Abbildung g von C auf P: die Endpunkte der Bögen $b_1, \ldots, b_4$ werden auf den Punkt 0, die Bögen selbst werden der Reihe nach proportional auf die Kreise U, V, U^{-1}, V^{-1} abgebildet; das Bild ist ein von 0 nach 0 zurückführender geschlossener Weg $C' = U V U^{-1} V^{-1}$. Jeder im Inneren von U liegende Punkt hat in bezug auf C' die Ordnung 0, da er von U einmal positiv, von U^{-1} einmal negativ, von V und V^{-1} gar nicht umlaufen wird; ebenso hat jeder im Inneren von V liegende Punkt in bezug auf C' die Ordnung 0, und dasselbe gilt für die außerhalb von U und V gelegenen Punkte von P. Jede Abbildung $g(E)$, zu der die Randabbildung $g(C) = C'$ gehört, hat also in allen (nicht auf C' gelegenen) Punkten den Grad 0. Es gibt daher sowohl Abbildungen g, bei denen -1, als solche, bei denen $+1$ nicht zu der Bildmenge gehört[29]. Wir behaupten aber: *Bei jeder Abbildung $g(E)$ mit $g(C) = C'$ gehört wenigstens einer der Punkte $-1, +1$ zur Bildmenge und zwar besteht die Originalmenge des Punktepaares $-1, +1$ sogar stets aus wenigstens zwei Punkten.*

Beweis[30]. Wir dürfen annehmen, daß einer der beiden Punkte keinen Originalpunkt hat, und infolge der Symmetrie ist es keine Einschränkung,

[30]) Ein zweiter Beweis läßt sich folgendermaßen führen: Hat einer der Punkte nur einen Originalpunkt, so läßt sich dieser auf Grund von Teil I, § 1, Satz IXa durch eine kleine Abänderung, die dem anderen der beiden Punkte keinen neuen Originalpunkt schafft, beseitigen, da der Grad in den beiden Punkten 0 ist. Man hat also nur zu zeigen, daß es nicht eintreten kann, daß beide Punkte keinen Originalpunkt haben. Dies würde aber, da dann eine Abbildung der von C berandeten Kreisscheibe in die in $+1$ und -1 punktierte Ebene vorläge, bedeuten, daß der Weg $U V U^{-1} V^{-1}$ in dieser zweimal punktierten Ebene zusammenziehbar wäre; das ist jedoch nicht der Fall, da deren Fundamentalgruppe bekanntlich die von U und V erzeugte freie Gruppe ist.

anzunehmen, daß -1 dieser Punkt sei. Dann ist zu zeigen, daß $+1$ wenigstens zwei Originalpunkte besitzt. g ist zugleich eine Abbildung von E auf die Zylinderfläche P', die durch Herausnahme von -1 aus P entsteht; fassen wir die Abbildung so auf, so bezeichnen wir sie mit g'. Die Behauptung wird bewiesen sein, wenn wir gezeigt haben, daß es bei g' in $+1$ zwei wesentliche Schichten gibt. (Bei g gibt es in jedem Punkt höchstens eine Schicht, da in P jeder geschlossene Weg zusammenziehbar ist.)

Da in E jeder geschlossene Weg zusammenziehbar ist, sind auch die durch g' gelieferten Bilder der geschlossenen Wege sämtlich zusammenziehbar, die Bildgruppe $\mathfrak{U}$ (siehe § 2, Definition I) besteht nur aus der Identität, und die zu g' gehörige Überlagerungsmannigfaltigkeit (Definition III) P^* von P' ist daher die „universelle“ Überlagerungsfläche, d. h. diejenige, die man erhält, wenn man in P' zwei von einem fest zugrunde gelegten Punkt ξ nach einem Punkt η laufende Wege w_1, w_2 dann und nur dann als denselben Punkt der Überlagerungsfläche auffaßt, wenn der geschlossene Weg $w_1 w_2^{-1}$ in P' zusammenziehbar ist. Nun ist die Zusammenziehbarkeit eines geschlossenen Weges v in P' gleichbedeutend damit, daß in P der Punkt -1 in bezug auf v die Ordnung 0 hat, daß sich also, wenn man in P ein Polarkoordinatensystem r, ψ mit -1 als Pol einführt, ψ auf v als eindeutige Funktion erklären läßt. Folglich liefern, wenn man in ξ einen Wert von ψ festgelegt hat, w_1 und w_2 dann und nur dann denselben Punkt von P^*, wenn die stetige Fortsetzung von ψ längs w_1 denselben Wert in η liefert wie die Fortsetzung längs w_2. Daraus folgt, daß man P^* darstellen kann als die durch $r > 0$ bestimmte Hälfte einer Ebene, in der r und ψ rechtwinklige kartesische Koordinaten sind. Die durch die Überlagerung gegebene Abbildung φ von P^* auf P' (siehe § 1) sowie die Abbildung g^* von E auf P^* (siehe Satz I) wird unmittelbar durch die Werte von r und ψ vermittelt.

Auf Grund des Satzes III und der Definitionen VII a, b haben wir zu zeigen, daß es unter den Punkten von P^*, die durch φ auf den Punkt $+1$ von P abgebildet werden, zwei gibt, in denen die Abbildung $g^*(E)$ von 0 verschiedene Grade oder, was dasselbe ist, die in bezug auf das Bild $C^* = g^*(C)$ von 0 verschiedene Ordnungen haben. Wählt man in dem r-ψ-Polarkoordinatensystem die Richtung der positiven x-Achse als $\psi = 0$, so sind die durch φ auf $+1$ abgebildeten Punkte diejenigen mit $r = 2$, $\psi = 2\pi m$, wobei m ganz ist. g^* sei überdies so normiert, daß ψ in dem (durch g' auf den Punkt 0 abgebildeten) Anfangspunkt des Bogens b_1 von C' den Wert 0 hat. Dann kann der Verlauf von $C^* = g^*(C)$ $= g^*(b_1 b_2 b_3 b_4) = g^*(b_1) g^*(b_2) g^*(b_3) g^*(b_4)$ angegeben werden: $g^*(b_1)$ ist die Strecke von $r = 1$, $\psi = 0$ nach $r = 1$, $\psi = 2\pi$; $g^*(b_2)$ ist eine

von dem letztgenannten Punkt in ihn zurücklaufende einfach geschlossene Kurve, die im Inneren des Streifens $2\pi - \frac{\pi}{2} < \psi < 2\pi + \frac{\pi}{2}$ liegt, symmetrisch zu der Geraden $\psi = 2\pi$ ist und mit ihr die Punkte $r = 1$, $r = 3$, und nur diese, gemeinsam hat: $g^*(b_3)$ ist die Strecke von $r = 1$, $\psi = 2\pi$ nach $r = 1$, $\psi = 0$; $g^*(b_4)$ ist die einfach geschlossene Kurve, die aus $g^*(b_2)^{-1}$ durch Translation längs der Strecke $g^*(b_3)$ entsteht. Dieser Verlauf von C^* zeigt, daß es unter den Punkten $r = 2$, $\psi = 2m\pi$ genau zwei gibt, die in bezug auf C^* von 0 verschiedene Ordnungen haben, nämlich die im Inneren der geschlossenen Kurven $g^*(b_2)$ bzw. $g^*(b_4)$ gelegenen Punkte mit $m = 1$ und $m = 0$; sie haben die Ordnungen $+1$ bzw. -1.

Dies bedeutet, daß es bei der Abbildung g' im Punkte $+1$ zwei wesentliche Schichten gibt; deren Beiträge sind je 1. Damit ist die Behauptung bewiesen.

Wenn wir statt der betrachteten Randabbildungen $g(C) = C'$ eine andere Randabbildung $g_1(C) = C_1'$ zugrunde legen, in die sich $g(C)$ stetig so überführen läßt, daß dabei in keinem Augenblick einer der Punkte -1, $+1$ auf dem Bilde von C liegt, so hat bei jeder Abbildung $g_1(E)$ mit der Randabbildung $g_1(C)$ das Punktepaar -1, $+1$ wieder wenigstens zwei Originalpunkte; denn in dem oben gegebenen Beweis ist nur der Weg $C^* = g^*(C)$ durch $C_1^* = g_1^*(C)$ zu ersetzen, wobei C_1^* aus C^* durch eine stetige Abänderung hervorgeht, während welcher niemals einer der Punkte $r = 2$, $\psi = 2m\pi$ auf der sich ändernden Kurve liegt, so daß die Ordnungen dieser Punkte in bezug auf die Kurve ungeändert bleiben, also für C_1^* dieselben sind wie für C^*.

Wir erweitern jetzt die anfangs betrachtete Abbildung $g(E)$ auf die ganze Ebene M, in der E liegt, durch die Festsetzung: das nicht im Inneren von E liegende, unendliche Stück jedes Strahles durch den Mittelpunkt von E wird auf den Punkt von C' abgebildet, der das Bild des auf dem Strahl liegenden Punktes von C ist. Da die Originalmenge jedes nicht auf C' liegenden Punktes von P ungeändert bleibt, bleibt der Grad überall 0. $g(M)$ ändern wir gleichmäßig stetig so ab, daß sie immer kompakt in jedem der Punkte -1, $+1$ bleibt. Ist $g_1(M)$ eine sich bei dieser Abänderung ergebende Abbildung, so hat, wie man leicht sieht, eine hinreichend große, E enthaltende, zu E konzentrische Kreisscheibe E_1 in M die Eigenschaft, daß das Bild des Randes C_1 von E_1 in keinem Augenblick des Überganges von g in g_1 einen der Punkte -1, $+1$ enthält. Folglich gibt es, da wir die im vorigen Absatz gemachte Bemerkung auf E_1 und C_1 statt auf E und C anwenden können, bei der Abbildung g_1 wenigstens zwei Originalpunkte von -1 und $+1$ in E_1.

Somit hat die Abbildung g der Ebene M auf die Ebene P die folgende Eigenschaft: sie hat überall, wo sie kompakt ist, den Grad 0; *sie ist in* -1 *und* $+1$ *kompakt; bei jeder Abbildung, die aus g durch eine solche gleichmäßig stetige Änderung hervorgeht, daß in jedem Augenblick die jeweilige Abbildung in den Punkten* -1, $+1$ *kompakt ist, haben diese beiden Punkte zusammen wenigstens zwei Originalpunkte.*

Dies ist das Ergebnis unserer Hilfsbetrachtung; mit seiner Hilfe konstruieren wir jetzt ein Beispiel einer Abbildung f einer Fläche M auf eine Fläche μ, so daß in einem Punkte ξ von μ, in dem f kompakt ist, $\sigma_\xi > s_\xi$ und $\alpha_\xi > a_\xi$ ist:

M sei die oben betrachtete gleichnamige Ebene. μ sei ein Zylinder; die universelle Überlagerungsfläche von μ werde durch die Ebene P derart dargestellt, daß zwei Punkte x, y und x', y' dann und nur dann zu demselben Punkt von μ gehören, wenn $x' = x + 2k$ (k ganz), $y' = y$ ist. ξ sei der Punkt von μ, der zu den Punkten $x = 2k + 1$, $y = 0$ von P gehört. Die durch die Überlagerung gegebene Abbildung von P auf μ bezeichnen wir mit Φ. Die Abbildung f von M auf μ sei durch $f(M) = \Phi g(M)$ definiert, wobei g dieselbe Bedeutung wie früher hat. Da jeder geschlossene Weg in M zusammenziehbar ist, so folgt — wie an einer ähnlichen Stelle während der obigen Hilfsbetrachtung —, daß die zu f gehörige Überlagerungsmannigfaltigkeit μ^* die universelle Überlagerungsfläche von μ ist. Wenn wir diese in der angegebenen Weise mit P identifizieren, so ist in unserer früheren Bezeichnungsweise $f^* = g$. Da g in allen Punkten $x = 2k + 1$, $y = 0$ kompakt ist und in jedem von ihnen den Grad 0 hat, ist f in ξ kompakt und besitzt dort keine wesentliche Schicht. Es ist also $s_\xi = a_\xi = 0$.

Die durch f bestimmte Klasse $\mathfrak{K}_\xi$ in bezug auf ξ ist die Gesamtheit aller Abbildungen, die man erhält, wenn man f gleichmäßig stetig so abändert, daß die Abbildungen immer kompakt in ξ bleiben. Einer solchen Änderung entspricht eine gleichmäßig stetige Änderung von $f^* = g$ (siehe Satz II), bei der die Abbildungen immer kompakt in allen Punkten $x = 2k + 1$, $y = 0$, also insbesondere in den Punkten -1 und $+1$, bleiben. Nach dem Ergebnis der Hilfsbetrachtung behält das Punktepaar -1, $+1$ dabei immer wenigstens zwei Originalpunkte, und diese sind Originalpunkte von ξ bei der abgeänderten Abbildung f. Folglich ist $\sigma_\xi \geqq 2$, und, da immer $\alpha_\xi \geqq \sigma_\xi$ ist, auch $\alpha_\xi \geqq 2$, also $\sigma_\xi > s_\xi$, $\alpha_\xi > a_\xi$, w. z. b. w. (Es ist übrigens leicht, durch Angabe einer speziellen Abbildung g zu zeigen, daß $\alpha_\xi = \sigma_\xi = 2$ ist.) Also ist bewiesen:

Satz XV. *Ist $n = 2$, so gibt es — im Gegensatz zu den von* 2 *verschiedenen Dimensionszahlen — Abbildungen, für die die im* § 4 *definierten Mindestzahlen σ_ξ und α_ξ größer sind als die Invarianten s_ξ bzw. a_ξ.*

Am Schluß des vorigen Paragraphen sahen wir, daß die Ausnahmestellung der Dimensionszahl 2 bezüglich der Zahlen α und a in Fortfall kommt, falls M geschlossen ist. Es entsteht die Frage, ob dies bezüglich der Zahlen σ und s ebenso ist, d. h. ob die für die anderen Dimensionszahlen richtige Gleichung $\sigma = s$ im Fall $n = 2$ wenigstens für alle Abbildungen von geschlossenen Flächen gilt. Diese Frage ist zu verneinen. Wir werden nämlich zeigen:

Satz XVa. *Zu jedem $j > 4$ gibt es eine Klasse von Abbildungen der geschlossenen orientierbaren Fläche vom Geschlecht 2 auf die geschlossene orientierbare Fläche vom Geschlecht 1 mit $\sigma > j = s$.*

Beweis. Sind μ, μ^* zwei geschlossene orientierbare Flächen vom Geschlecht 1, so läßt sich μ^* durch eine Abbildung φ so auf μ abbilden, daß μ^* eine j-blätterige unverzweigte Überlagerungsfläche von μ wird. f^* sei folgende Abbildung der geschlossenen orientierbaren Fläche M vom Geschlecht 2 auf μ^*: durch eine geeignete einfach geschlossene Kurve C wird M in zwei Hälften zerlegt, deren jede eine einmal berandete Fläche vom Geschlecht 1 ist; jede der beiden Hälften wird so auf μ^* abgebildet, daß die Randkurve in einen einzigen Punkt ζ von μ^* übergeht und daß die Abbildung im übrigen auf jeder der Hälften eineindeutig vom Grade $+1$ ist; dann hat f^* den Grad $+2$ und, da die Originalmenge von ζ das Kontinuum C ist, nach Satz IX (§ 3) den Index 1, und nach Satz X ist jeder geschlossene Weg auf μ^* dem Bilde eines geschlossenen Weges auf M äquivalent. Hieraus folgt für die Abbildung $f = \varphi f^*$ von M auf μ, daß μ^* die zu ihr gehörige Überlagerungsmannigfaltigkeit (Definition III, § 2), die Zerlegung in f^* und φ die durch Satz I gekennzeichnete Zerlegung und mithin j der Index von f ist. Wenn es nun in der Klasse von f eine Abbildung f_1 gibt, bei der ein Punkt ξ von μ nur j Originalpunkte hat, so hat bei der f_1 entsprechenden Abbildung f_1^* von M auf μ^* jeder der j Punkte von μ^*, die durch φ auf ξ abgebildet werden, nur einen Originalpunkt. Daher ist die Behauptung bewiesen, sobald gezeigt ist, daß es bei jeder Abbildung von M auf μ^* vom Grade 2 höchstens 4 Punkte auf μ^* gibt, die nur je einen Originalpunkt haben; somit ist der Satz XVa zurückgeführt auf

Satz XVI. *Bei jeder Abbildung der geschlossenen orientierbaren Fläche F_p vom Geschlecht p auf die geschlossene orientierbare Fläche F_q vom Geschlecht q, deren Grad einen Betrag > 1 hat, ist die Anzahl der Punkte auf F_q, die nur je einen Originalpunkt haben, höchstens $2p + 2 - 2q$.*

Der Beweis wird nicht wie die übrigen in dieser Arbeit angestellten Betrachtungen mit dem auf der stetigen Deformierbarkeit beruhenden Begriff der „Äquivalenz“ geschlossener Wege, sondern mit dem Begriff der „Homo-

logie" arbeiten[31]), und wir schicken einen in dieser Richtung liegenden einfachen Hilfssatz voraus.

Hilfssatz. Die offene Fläche F_q', die aus der geschlossenen orientierbaren Fläche F_q vom Geschlecht q durch Herausnahme von s (>0) Punkten $y_1, y_2, \ldots, y_s$ entsteht, besitzt eine Homologiebasis von $2q+s-1$ Elementen (d. h. es gibt auf F_q' $2q+s-1$ geschlossene Kurven derart, daß jede geschlossene Kurve auf F_q' einer eindeutig bestimmten linearen Verbindung von ihnen homolog ist).

Beweis des Hilfssatzes. $A_1, A_2, \ldots, A_{2q}$ seien geschlossene Kurven auf F_q', die eine Homologiebasis von F_q bilden; $B_1, B_2, \ldots, B_s$ seien kleine Kreise um die Punkte $y_1, y_2, \ldots, y_s$. Ist C irgendeine geschlossene Kurve auf F_q', so ist C auf F_q einer linearen Verbindung der A_i homolog:

$$C \sim \sum_{i=1}^{2q} a_i A_i \quad (\text{auf } F_q), \tag{1}$$

d. h. es gibt einen auf F_q liegenden Flächenkomplex K mit dem Rand $C - \sum a_i A_i$. Schneidet man aus F_q so kleine Löcher $L_1, L_2, \ldots, L_s$ um $y_1, y_2, \ldots, y_s$ aus, daß kein Randpunkt von K in einem solchen Loch liegt, so entsteht aus K ein Komplex K', der F_q' angehört und dessen Rand außer aus dem Rande von K aus einer linearen Verbindung der Ränder der Löcher besteht; da diese Ränder auf F_q' den Kreisen B_j homolog sind, ist also der Rand von K auf F_q' einer linearen Verbindung der B_j homolog, mithin ist

$$C \sim \sum_{i=1}^{2q} a_i A_i + \sum_{j=1}^{s} b_j B_j \quad (\text{auf } F_q'). \tag{2}$$

Die A_i und B_j zusammen erzeugen also die Gruppe der Homologieklassen von F_q'. Sie sind aber nicht voneinander unabhängig; denn die Summe der B_j bildet den Rand des Teiles von F_q', der aus F_q durch Herausnahme der durch die B_j begrenzten, die Punkte y_j enthaltenden kleinen Kreisscheiben entsteht; d. h. es ist

$$\sum_{j=1}^{s} B_j \sim 0 \quad (\text{auf } F_q'). \tag{3}$$

Dies ist aber die einzige zwischen den A_i und B_j auf F_q' bestehende Relation; denn wenn

$$\sum_{i=1}^{2q} u_i A_i + \sum_{j=1}^{s} v_j B_j \sim 0 \quad (\text{auf } F_q') \tag{4}$$

irgendeine Relation ist, so folgt, da (4) erst recht auf F_q gilt, aus

$$B_j \sim 0 \quad (\text{auf } F_q), \tag{5}$$

[31]) Zur Orientierung über die mit der „Homologie" zusammenhängenden Begriffe und Sätze vgl. man: J. W. Alexander, Combinatorial Analysis Situs. Transact. Am. Math. Soc. 28 (1926).

daß

$$\sum_{i=1}^{2q} u_i A_i \sim 0 \quad (\text{auf } F_q),$$

also

(6) $$u_i = 0 \qquad (i = 1, 2, \ldots, 2q),$$

mithin nach (4)

(7) $$\sum_{j=1}^{s} v_j B_j \sim 0 \quad (\text{auf } F_q')$$

ist. Elimination von B_s aus (3) und Einsetzen in (7) liefert

(8) $$\sum_{j=1}^{s-1} (v_j - v_s) B_j \sim 0 \quad (\text{auf } F_q').$$

Dies bedeutet die Existenz eines zweidimensionalen Teilkomplexes K_1 von F_q', der von $\sum (v_j - v_s) B_j$ berandet wird; fügt man zu ihm für $i = 1, 2, \ldots, s-1$ das $(v_s - v_j)$-mal genommene Innengebiet von B_j (d. h. das von B_j begrenzte, y_j enthaltende Gebiet) hinzu, so entsteht ein *geschlossener* Komplex K_2, der in einem *echten* Teil der geschlossenen Fläche F_q liegt, da er den Punkt y_s nicht enthält. Ein solcher Komplex ist aber identisch 0; und da die $(v_s - v_j)$-mal genommene Umgebung von y_j ein Stück von K_1 ist, muß daher $v_s = v_j$ für alle j sein. Hieraus und aus (6) folgt, daß (4) die Gestalt

$$v_s \cdot \sum_{j=1}^{s} B_j \sim 0$$

hat, also eine Folge von (3) ist.

Man kann also jede geschlossene Kurve auf F_q' in der Gestalt (2) darstellen, und zwischen den A_i und B_j besteht nur die Relation (3). Damit ist der Hilfssatz bewiesen; denn man kann als Basis zum Beispiel die Kurven $A_1, A_2, \ldots, A_{2q}, B_1, \ldots, B_{s-1}$ wählen.

Beweis von Satz XVI. F_p sei durch f auf F_q mit dem Grade c abgebildet, und es sei $|c| > 1$. $y_1, y_2, \ldots, y_r$ seien Punkte auf F_q, die nur je einen Originalpunkt haben; ob es noch mehr solche Punkte gibt, ist dabei gleichgültig. Die ihnen entsprechenden Originalpunkte seien $x_1, x_2, \ldots, x_r$. Die offenen Flächen, die durch Herausnahme der Punkte $x_1, x_2, \ldots, x_{r-1}$ bzw. $y_1, y_2, \ldots, y_{r-1}$ aus F_p bzw. F_q entstehen, seien mit F_p', F_q' bezeichnet. Dann ist $f(F_p')$ eine überall kompakte Abbildung von F_p' auf F_q', bei der der Punkt y_r nur einen einzigen Originalpunkt hat, bei der daher nach Satz X (§ 3) jeder geschlossene Weg auf F_q' dem Bild eines geschlossenen Weges auf F_p' äquivalent, also a fortiori homolog ist.

Nach dem Hilfssatz gibt es auf F_q' eine Basis $C_1, C_2, \ldots, C_{2q+r-2}$. $Z_1, Z_2, \ldots, Z_{2q+r-2}$ seien geschlossene Wege auf F_p' deren Bilder den

Wegen C_i homolog sind:

$$(9) \qquad f(Z_i) \sim C_i \quad (\text{auf } F'_q;\ i = 1, 2, \ldots, 2q + r - 2).$$

Wir behaupten, daß die Z_i unabhängig voneinander bezüglich Homologien nicht nur auf F'_p, sondern sogar auf F_p sind.

Ist

$$(10) \qquad \sum a_i Z_i \sim 0 \quad (\text{auf } F_p),$$

so können wir, da es auf F_p keine „Nullteiler" gibt, d. h. da auf F_p ein Vielfaches einer geschlossenen Kurve nur dann ~ 0 ist, wenn die einfache Kurve ~ 0 ist, die linke Seite von (10) durch den größten gemeinsamen Teiler der a_i dividieren. Wir dürfen also von vornherein annehmen, daß die a_i entweder zueinander teilerfremd oder sämtlich 0 sind. (10) besagt, daß es auf F_p einen zweidimensionalen Komplex K gibt, der von $\sum a_i Z_i$ berandet wird. Schneiden wir um die Punkte $x_1, x_2, \ldots, x_{r-1}$ herum so kleine Elemente $L_1, L_2, \ldots, L_{r-1}$ aus F_p aus, daß kein Randpunkt von K (also kein Punkt einer Kurve Z_i) in einem L_j und daß jedes Bild $f(L_j)$ in einer euklidischen Umgebung von y_j liegt, so entsteht aus K ein Komplex K', der außer von dem Rande von K noch von einer linearen Verbindung der Ränder $Y_1, Y_2, \ldots, Y_{r-1}$ der L_j berandet wird; da K' in F'_p liegt, ist daher

$$(11) \qquad \sum a_i Z_i + \sum b_j Y_j \sim 0 \quad (\text{auf } F'_p).$$

Die Abbildung $f(L_j)$ hat, da x_j der einzige Originalpunkt von y_j ist, im Punkt y_j den Grad c; die Ordnung von y_j in bezug auf die Bildkurve $f(Y_j)$, d. h. die Anzahl der Umläufe von $f(Y_j)$ um y_j, ist daher c. Wenn also $\overline{Y}_j$ einen, den Punkt y_j einmal umlaufenden kleinen Kreis auf F_q bezeichnet, so ist

$$(12) \qquad f(Y_j) \sim c\,\overline{Y}_j \quad (\text{auf } F'_q).$$

Aus (11), (9), (12) folgt

$$(13) \qquad \sum a_i C_i + c \sum b_j \overline{Y}_j \sim 0 \quad (\text{auf } F'_q).$$

$\sum a_i C_i$ ist also auf F'_q einer c-mal genommenen geschlossenen Kurve homolog; da sich jede geschlossene Kurve eindeutig als Verbindung der C_i darstellen läßt, folgt daraus, daß alle a_i durch c teilbar sind. Da $|c| > 1$ ist, sind sie daher nicht teilerfremd, also, wie wir oben sahen, sämtlich 0. Dies bedeutet im Hinblick auf ihre Definition (10), daß die Z_i $(i = 1, 2, \ldots, 2q + r - 2)$ auf F_p voneinander unabhängig sind. Ihre Anzahl kann also höchstens $2p$ sein, d. h. es ist

$$2p \geqq 2q + r - 2, \qquad r \leqq 2p + 2 - 2q.$$

Damit sind der Satz XVI und der oben auf diesen zurückgeführte Satz XVa bewiesen.

Anhang I.

Über die Punktmenge, in der eine Abbildung eineindeutig ist.

Die soeben durchgeführte Abschätzung der Anzahl derjenigen Punkte bei einer Flächenabbildung, die nur je einen einzigen Originalpunkt haben, läßt sich — wenigstens unter gewissen Einschränkungen — auf Abbildungen n-dimensionaler Mannigfaltigkeiten übertragen. An Stelle der *Anzahl* der Punkte, die ja die 0-dimensionale Bettische Zahl des Komplexes dieser Punkte ist, tritt dabei die $(n-2)$-*dimensionale Bettische Zahl*[31]), und diese Betrachtung weist daher vielleicht einen Weg, auf dem man den Fall $n=2$ von seiner Ausnahmestellung, die in dieser Arbeit eine so große Rolle spielt, befreien und ihn in eine allgemeine Theorie einordnen kann: man wird versuchen müssen, nicht nur die Komponentenzahlen der Originalmengen einzelner Punkte oder der glatten Originalmengen einzelner Gebiete, sondern auch die höheren Zusammenhangszahlen dieser Mengen zu untersuchen.

Die erwähnten Einschränkungen — die übrigens kaum wesentlich sein dürften — bestehen darin, daß wir nur solche Mannigfaltigkeiten und Abbildungen betrachten, die den Methoden der kombinatorischen Topologie ohne weiteres zugänglich sind. Wir werden uns also auf *triangulierbare* geschlossene Mannigfaltigkeiten beschränken, während den bisherigen Untersuchungen die (wenigstens begrifflich) weitere Gesamtheit der „topologischen" Mannigfaltigkeiten zugrunde lag, und wir werden von den Abbildungen voraussetzen, daß sie „*simplizial*" sind, d. h. daß sie den Eckpunkten eines Simplexes von M immer Eckpunkte eines (evtl. niedriger-dimensionalen) Simplexes von μ zuordnen und daß sie im Inneren der Simplexe von M die durch die Eckpunktzuordnung eindeutig bestimmten baryzentrischen Abbildungen sind.

Der Grad c einer simplizialen Abbildung von M auf μ ist die Anzahl der auf ein festes n-dimensionales Simplex von μ im positiven Sinne abgebildeten n-dimensionalen Simplexe von M vermindert um die Anzahl der auf dasselbe Simplex im negativen Sinne abgebildeten n-dimensionalen Simplexe von M. Die Punkte von μ, die nur je einen Originalpunkt haben, bilden einen Komplex $\overline{X}$, der, wenn $|c|>1$ ist, höchstens $(n-2)$-dimensional ist; denn würde er ein $(n-1)$-dimensionales Simplex τ^{n-1} enthalten, so kämen als Originalsimplexe für die beiden an τ^{n-1} anstoßenden n-dimensionalen Simplexe nur die beiden an das Originalsimplex t^{n-1} von τ^{n-1} anstoßenden Simplexe in Betracht, während doch jedes n-dimensionale Simplex von μ wenigstens $|c|$ Originalsimplexe besitzt. Die $(n-2)$-dimensionale, also die höchste nicht triviälerweise verschwindende, Bettische Zahl von $\overline{X}$, und damit auch von dem mit $\overline{X}$ homöomorphen Original-

komplex X von $\overline{X}$ ist, wie wir zeigen werden, durch die Bettischen Zahlen von M und μ nach oben beschränkt.

Wir bezeichnen die i-te Bettische Zahl eines Komplexes K stets mit $p^i(K)$.

Satz XVII. *M sei auf μ simplizial mit dem Grade c abgebildet; $|c|$ sei > 1 und, falls M 1-dimensionale Torsionskoeffizienten besitzt, zu diesen teilerfremd; X und $\overline{X}$ seien die Komplexe in M bzw. μ, auf denen die Abbildung eineindeutig ist. Ist dann X' ein echter Teilkomplex von X, so ist*

$$p^{n-2}(X') \leqq p^1(M) - p^1(\mu) + p^2(\mu).$$

Bevor wir den Satz beweisen, werden wir einige Folgerungen ziehen. Zunächst ergibt sich, da nur dann $c \neq 0$ sein kann, wenn $p^2(M) \geqq p^2(\mu)$ ist[32]), unmittelbar eine Beschränkung von $p^{n-2}(X')$ mittels der Invarianten von M allein:

Satz XVIIa. *Unter den Voraussetzungen von Satz* XVII *ist*

$$p^{n-2}(X') \leqq p^1(M) + p^2(M).$$

Hierin ist z. B. die Tatsache enthalten, daß, falls bei einer Abbildung der 3-dimensionalen Sphäre mit $|c| > 1$ der Komplex X ein einfach geschlossenes Polygon enthält, er mit diesem identisch ist, da die erste Bettische Zahl eines einfach geschlossenen Polygons 1, die eines *echten* Teils von X nach XVIIa aber 0 ist. Dasselbe gilt für die Abbildungen des 3-dimensionalen projektiven Raumes, vorausgesetzt, daß nicht nur $|c| < 1$, sondern auch c ungerade ist, da der projektive Raum den 1-dimensionalen Torsionskoeffizienten 2 besitzt. Daß hierbei die Voraussetzung der Ungeradheit von c für die Gültigkeit von XVII und XVIIa wirklich notwendig ist, zeigt folgendes Beispiel: Der euklidische x_1-x_2-x_3-Raum sei der Abbildung auf sich unterworfen, die, wenn man statt x_1, x_2 Polarkoordinaten r, φ einführt, durch

$$r' = r, \quad \varphi' = 2\varphi, \quad x_3' = x_3$$

gegeben ist und die offenbar den Grad 2 hat; die Transformation von x_1, x_2, x_3 ist

$$x_1' = \frac{x_1^2 - x_2^2}{\sqrt{x_1^2 + x_2^2}}, \qquad x = \frac{2x_1 x_2}{\sqrt{x_1^2 + x_2^2}}, \qquad x_3' = x_3,$$

und sie läßt sich durch $x_4' = x_4$ zu einer Abbildung des projektiven $(x_1 : x_2 : x_3 : x_4)$-Raumes auf sich erweitern. Bei dieser hat sowohl jeder Punkt mit $x_1 = x_2 = 0$ als jeder Punkt mit $x_3 = x_4 = 0$ nur einen Original-

---

[32]) H. Hopf, On some properties of one-valued transformations of manifolds, Satz Ia. Proc. of the Nat. Acad. of Sciences U. S. A. **14** (1928). — Eine ausführliche Darstellung erscheint demnächst unter dem Titel „Zur Algebra der Abbildungen von Mannigfaltigkeiten" im Journ. f. d. reine u. angew. Math.

punkt, X enthält also *zwei* geschlossene projektive Geraden (und ist übrigens mit diesen identisch) — im Gegensatz zu den Behauptungen der beiden Sätze[33]).

Auch die $(n-2)$-te Bettische Zahl von X selbst läßt sich abschätzen: X' entstehe durch Entfernung eines $(n-2)$-dimensionalen Simplex T aus X (den Fall, daß X weniger als $(n-2)$-dimensional ist, brauchen wir nicht zu berücksichtigen, da dann $p^{n-2}(X)=0$ ist). Beim Übergang von X' zu X, d. h. beim Einsetzen von T in X' sind zwei Fälle möglich: entweder entsteht kein neuer $(n-2)$-dimensionaler Zyklus; dann ist $p^{n-2}(X)=p^{n-2}(X')$; oder es entsteht wenigstens ein neuer Zyklus Z; er enthalte T a-fach genommen; ist dann Z' ein zweiter Zyklus in X, der T enthält, und zwar b-mal genommen, so ist $aZ'-bZ$ ein Zyklus in X', d. h. Z und Z' sind in X' voneinander linear abhängig, und der Rang der Gruppe der $(n-2)$-dimensionalen Zyklen hat sich somit beim Übergang von X' zu X nur um 1 vermehrt, in diesem Falle ist also $p^{n-2}(X)=p^{n-2}(X')+1$. In jedem Fall ist $p^{n-2}(X)\leqq p^{n-2}(X')+1$. Mithin folgt aus XVII und XVIIa:

Satz XVIIb. *Unter den Voraussetzungen von* XVII *ist*

$$p^{n-2}(X)\leqq 1+p^1(M)-p^1(\mu)+p^2(\mu)\leqq 1+p^1(M)+p^2(M).$$

Für $n=2$ ist die erste dieser Ungleichungen der Satz XVI, da dann $p^2(\mu)=p^2(F_q)=1$ und $p^0(X)$ die Anzahl der Punkte ist, aus denen X besteht.

Der Beweis von Satz XVII läuft dem Beweis des Satzes XVI ganz parallel. Wir haben zunächst eine Tatsache festzustellen, die dem dem Beweis von XVI vorangeschickten Hilfssatz entspricht. Sie beruht auf einer von Pontrjagin entdeckten Verallgemeinerung des Alexanderschen Dualitätssatzes, die folgendermaßen lautet[34]): *Ist K ein in μ liegender Komplex, so werden die in K gelegenen i-dimensionalen Zyklen, die ~ 0 in μ sind, in Klassen bezüglich Homologien in K eingeteilt*; *der Rang der Gruppe dieser Klassen heiße $r^i(K)$*; *analog seien für die Komplementär-*

[33]) Um den Gegensatz zu dem Wortlaut dieser Sätze vollständig zu machen, hat man die im Text angegebene Abbildung noch durch eine simpliziale Abbildung mit denselben Eigenschaften zu ersetzen, was aber infolge des analytischen Charakters der Abbildung keine Schwierigkeit macht.

[34]) Pontrjagin, Zum Alexanderschen Dualitätssatz, Zweite Mitteilung, Satz Ia. Nachr. d. Ges. d. Wiss. zu Göttingen, Math.-Phys. Klasse (1927). Dort ist der Satz für „Homologien mod 2" bewiesen; der Beweis für gewöhnliche Homologien ist mir aus einer Note von Pontrjagin bekannt, die demnächst veröffentlicht werden dürfte. — In engem Zusammenhang damit stehen die folgenden Arbeiten: van Kampen, Eine Verallgemeinerung des Alexanderschen Dualitätssatzes. Koninkl. Akad. van Wetenschappen te Amsterdam Proc. **31** (1928), und: Die kombinatorische Topologie und die Dualitätssätze, Diss. Leiden 1929. — Lefschetz, Closed point sets on a manifold. Annals of Math. **29** (1928), sowie die dort zitierten Noten in den Proc. Nat. Acad. (1927).

menge von K die Zahlen $r^i(\mu - K)$ *erklärt. Dann ist*

$$r^i(\mu - K) = r^{n-1-i}(K),$$

also insbesondere

(1) $$r^1(\mu - K) = r^{n-2}(K).$$

Nun bestehen einfache Zusammenhänge zwischen diesen Zahlen r und den Bettischen Zahlen von K, $\mu - K$ und μ. Bezeichnet $P^i(K)$ die Gruppe der Klassen, in die *alle* i-dimensionalen Zyklen in K bezüglich Homologien *in* K zerfallen, deren Rang also $p^i(K)$ ist, $R^i(K)$ die oben betrachtete Untergruppe vom Range $r^i(K)$, so ist der Rang $p^i(K) - r^i(K)$ der Faktorgruppe $\frac{P^i(K)}{R^i(K)}$ höchstens gleich $p^i(\mu)$; denn von je $1 + p^i(\mu)$ Elementen der Faktorgruppe gibt es eine lineare Verbindung, die ~ 0 in μ, die also in $R^i(K)$ enthalten ist, je $1 + p^i(\mu)$ Elemente der Faktorgruppe sind also voneinander linear abhängig. Es ist mithin in der Tat

(2) $$p^i(K) - r^i(K) \leqq p^i(\mu),$$

also insbesondere

(2′) $$r^{n-2}(K) \geqq p^{n-2}(K) - p^{n-2}(\mu).$$

Analog ist

(3) $$p^i(\mu - K) - r^i(\mu - K) \leqq p^i(\mu),$$

also insbesondere

(3′) $$p^1(\mu - K) \leqq r^1(\mu - K) + p^1(\mu).$$

Es sei jetzt K höchstens $(n-2)$-dimensional. Dann steht in (3′) immer das Gleichheitszeichen; denn da es zu jedem 1-dimensionalen Zyklus in μ einen beliebig benachbarten, also homologen, gibt, der fremd zu K ist, also in $\mu - K$ liegt, gibt es in $\mu - K$ $p^1(\mu)$ Zyklen, die nicht homolog 0 in μ und die untereinander unabhängig bezüglich Homologien in μ, also erst recht bezüglich Homologien in $\mu - K$ sind; die ihnen entsprechenden Elemente der Faktorgruppe $\frac{P^1(\mu - K)}{R^1(\mu - K)}$ sind voneinander unabhängig, weil ja sonst eine lineare Verbindung von ihnen zu $R^1(\mu - K)$ gehörte, also ~ 0 in μ wäre. Folglich ist der Rang der Faktorgruppe nicht nur wie im allgemeinen Fall $\leqq p^1(\mu)$, sondern $= p^1(\mu)$, d. h. es ist in der Tat

(3″) $$p^1(\mu - K) = r^1(\mu - K) + p^1(\mu).$$

Ersetzt man hierin $r^1(\mu - K)$ aus (1), so folgt aus (2′):

(4) $$p^1(\mu - K) \geqq p^{n-2}(K) - p^{n-2}(\mu) + p^1(\mu),$$

wofür man infolge des Poincaréschen Dualitätsgesetzes für die Bettischen Zahlen einer Mannigfaltigkeit auch schreiben kann:

(4′) $$p^1(\mu - K) \geqq p^{n-2}(K) - p^2(\mu) + p^1(\mu).$$

Diese Ungleichung, die unter der Voraussetzung gilt, daß K höchstens $(n-2)$-dimensional ist, wird beim Beweise des Satzes XVII dieselbe Rolle spielen, wie der „Hilfssatz" beim Beweis von Satz XVI.

Wir betrachten nun eine Abbildung f, die die Voraussetzungen von Satz XVII erfüllt; X' sei ein echter Teilkomplex des in dem Satz genannten Komplexes X, $\overline{X}'$ das Bild von X'. Wir wählen in $\mu - \overline{X}'$ eine 1-dimensionale Homologiebasis; sie besteht aus $p^1(\mu - \overline{X}')$ geschlossenen Kurven $C_1, C_2, \ldots$ und, falls 1-dimensionale Torsionskoeffizienten $\tau_1, \tau_2, \ldots$ vorhanden sind, noch aus weiteren geschlossenen Kurven $D_1, D_2, \ldots$ derart, daß jeder in $\mu - \overline{X}'$ gelegene 1-dimensionale Zyklus C eine und nur eine Homologie

$$C \sim \sum a_i C_i + \sum b_j D_j, \quad 0 \leqq b_j < \tau_j, \quad (\text{in } \mu - \overline{X}')$$

erfüllt. Aus der Einzigkeit dieser Homologie folgt insbesondere, daß nur dann $C \sim c C'$ in $\mu - \overline{X}'$ sein kann, wenn alle a_i durch c teilbar sind.

$M - X'$ wird durch f so auf $\mu - \overline{X}'$ abgebildet, daß diese Abbildung überall kompakt ist. Dabei hat jeder Punkt von $\overline{X} - \overline{X}'$ nur einen Originalpunkt. Folglich ist nach Satz X jeder geschlossene Weg in $\mu - \overline{X}'$ dem Bilde eines geschlossenen Weges aus $M - X'$ in $\mu - \overline{X}'$ äquivalent, also erst recht homolog. $Z_1, Z_2, \ldots$ seien geschlossene Wege in $M - X'$, so daß

$$(5) \qquad f(Z_i) \sim C_i \quad (\text{in } \mu - \overline{X}';\ i = 1, 2, \ldots, p^1(\mu - \overline{X}'))$$

ist. Wir behaupten, daß diese Z_i unabhängig bezüglich Homologien nicht nur in $M - X'$, sondern sogar in M sind.

Dem Beweise schicken wir die Bemerkung voran, daß aus

$$(6) \qquad cZ \sim 0 \quad (\text{in } M)$$

stets

$$(7) \qquad Z \sim 0 \quad (\text{in } M)$$

folgt, wobei Z irgendein 1-dimensionaler Zyklus ist. Bilden nämlich die Zyklen $A_1, A_2, \ldots, A_{p^1(M)}, B_1, B_2, \ldots, B_{q^1(M)}$, wobei $p^1(M)$ die 1-dimensionale Bettische Zahl, $q^1(M)$ die Anzahl der 1-dimensionalen Torsionskoeffizienten $t_1, t_2, \ldots$ von M ist, eine Basis in M, und ist

$$Z \sim \sum u_i A_i + \sum v_j B_j, \quad 0 \leqq v_j < t_j \quad (\text{in } M),$$

so ist nach (6)

$$\sum c u_i A_i - \sum c v_j B_j \sim 0 \quad (\text{in } M),$$

also sind einerseits alle $u_i = 0$, andererseits sind alle $c v_j$ durch die entsprechenden t_j teilbar, was wegen der vorausgesetzten Teilerfremdheit von c und t_j nur möglich ist, wenn auch alle $v_j = 0$ sind, wenn also in der Tat (7) gilt.

Wenn nun

$$(8)\qquad \sum a_i Z_i \sim 0 \quad (\text{in } M)$$

ist, so können wir, falls alle a_i durch c teilbar sind, nach der eben gemachten Bemerkung die linke Seite von (8) durch c dividieren, und diese Division, falls nicht alle $a_i = 0$ sind, so oft wiederholen, bis die Koeffizienten in (8) nicht mehr c als gemeinsamen Teiler haben. Wir dürfen daher annehmen, daß von vornherein die a_i entweder sämtlich gleich 0 sind oder nicht den gemeinsamen Teiler c haben.

(8) bedeutet, daß ein 2-dimensionaler Komplex K in M existiert, der von $\sum a_i Z_i$ berandet wird. Ohne Beschränkung der Allgemeinheit läßt sich annehmen, daß sich K und X in „allgemeiner Lage" zueinander befinden, daß sie sich also in endlich vielen Punkten schneiden (da X als $(n-2)$-dimensional vorausgesetzt werden darf, weil sonst unsere Behauptung XVII von selbst erfüllt ist). Es seien also $x_1, x_2, \ldots, x_r$ die Schnittpunkte von X und K; dabei dürfen wir infolge der „allgemeinen Lage" noch annehmen, daß jeder Punkt x_j innerer Punkt eines $(n-2)$-dimensionalen Simplexes T_j^{n-2} von X und eines Dreiecks von K ist. Um x_j herum schneiden wir ein kleines Loch L_j aus K aus; der dadurch entstehende Komplex K' liegt ganz in $M-X$, also erst recht in $M-X'$, und wird außer von dem Rande von K noch von einer linearen Verbindung der Ränder der L_j berandet; nun ist der Rand von L_j in $M-X'$ homolog dem geschlossenen Polygon Y_j, welches von den zu T_j^{n-2} fremden Kanten derjenigen n-dimensionalen Simplexe $T_1^n, T_2^n, \ldots, T_s^n$ gebildet wird, auf denen T_j^{n-2} liegt. Es ist also

$$(9)\qquad \sum a_i Z_i + \sum b_j Y_j \sim 0 \quad (\text{in } M-X').$$

Da T_j^{n-2} zu X gehört, ist $f(T_j^{n-2}) = \overline{T}_j^{n-2}$ ein ebenfalls $(n-2)$-dimensionales Simplex und besitzt außer T_j^{n-2} kein weiteres Originalsimplex; mithin besitzen auch diejenigen n-dimensionalen Simplexe von μ, auf denen $\overline{T}_j^{n-2}$ liegt, als Originalsimplexe nur die $T_1^n, T_2^n, \ldots, T_s^n$, und daher hat die Abbildung des „Sternes" $\sum_{i=1}^{s} T_i^n$ von T_j^{n-2} auf den „Stern" von $\overline{T}_j^{n-2}$ den Grad c; dabei wird jede nicht an $\overline{T}_j^{n-2}$ anstoßende Kante des Sternes von $\overline{T}_j^{n-2}$ von zu Y_j gehörigen Kanten im algebraischen Sinne c-mal bedeckt, und andererseits wird jede zu Y_j gehörige Kante auf eine nicht an $\overline{T}_j^{n-2}$ anstoßende (evtl. in einen Punkt ausgeartete) Kante des Sternes von $\overline{T}_j^{n-2}$ abgebildet. $f(Y_j)$ ist also das c-mal genommene, aus den genannten Kanten des Sternes von $\overline{T}_j^{n-2}$ gebildete geschlossene Polygon η_j:

$$(10)\qquad f(Y_j) \sim c\,\eta_j \quad (\text{in } \mu - \overline{X}').$$

Aus (9), (5), (10) folgt

$$(11)\qquad \sum a_i C_i \sim c\,C' \quad (\text{in } \mu - \overline{X}'),$$

wobei $-\sum b_j \eta_j = C'$ gesetzt ist. Daraus folgt weiter, wie wir oben sahen, daß alle a_i durch c_i teilbar sind, und dies bedeutet, wie wir ebenfalls schon sahen, daß alle $a_i = 0$, daß also die Z_i unabhängig in M sind. Ihre Anzahl ist daher höchstens $p^1(M)$, d. h. es ist

$$p^1(M) \geqq p^1(\mu - \overline{X}^1), \tag{12}$$

also auf Grund von $(4')$

$$p^1(M) \geqq p^{n-2}(\overline{X}') - p^2(\mu) + p^1(\mu), \tag{13}$$

womit, da $p^i(\overline{X}') = p^i(X')$ ist, die Behauptung des Satzes XVII bewiesen ist.

Anhang II.

Über die Windungspunkte einer Flächenabbildung.

Hat man, wie es im Satz XVI und im Anhang I geschehen ist, bei einer Abbildung, deren Grad einen Betrag > 1 besitzt, diejenigen Punkte der Bildmannigfaltigkeit betrachtet, die nur je einen Originalpunkt haben, so liegt folgende Verallgemeinerung der Fragestellung nahe. Wir definieren: *Bei einer Abbildung mit einem Grade c, dessen Betrag > 1 ist, heißt ein Punkt der Bildmannigfaltigkeit ein „Windungspunkt“, wenn er weniger als $|c|$ Originalpunkte besitzt; ist deren Anzahl b, so heißt $|c| - b$ die „Ordnung“ des Windungspunktes.*

Für die Anzahlen der Windungspunkte der verschiedenen Ordnungen, die bei Flächenabbildungen auftreten, existiert eine Schranke, die eine Verallgemeinerung und Verschärfung des Satzes XVI liefert:

Satz XVIII. *Bei jeder Abbildung einer geschlossenen orientierbaren Fläche auf eine andere geschlossene orientierbare Fläche gibt es höchstens endlich viele Windungspunkte. Ist w_r die Anzahl der Windungspunkte der Ordnung r, so ist*

$$\sum_{r=1}^{|c|-1} r\, w_r \leqq (2p-2) - |c|(2q-2),$$

wobei p das Geschlecht der Originalfläche, q das Geschlecht der Bildfläche, c der Grad ist.

Der Beweis besteht in einer einfachen Anwendung des folgenden Satzes von H. Kneser[6]): *Ist die geschlossene orientierbare Fläche vom Geschlecht P auf die geschlossene orientierbare Fläche vom Geschlecht $Q \neq 0$ mit dem Grade $c \neq 0$ abgebildet, so ist*

$$(P-1) - |c|(Q-1) \geqq 0.$$

Um unsere Behauptung auf diesen Satz zurückzuführen, nehmen wir mit der gegebenen Abbildung f der Fläche F_p vom Geschlecht p auf die

Fläche F_q vom Geschlecht q eine unwesentliche Abänderung in der Nähe von Windungspunkten vor. Es seien für $r = 1, 2, \ldots, |c| - 1$ wenigstens je v_r Windungspunkte der Ordnung r vorhanden: $\xi_1^r, \xi_2^r, \ldots, \xi_{v_r}^r$, wobei es gleichgültig ist, ob dies alle Windungspunkte sind. Um jeden Punkt ξ_i^r bestimmen wir ein Element ω_i^r, so daß diese Elemente zueinander fremd sind. $x_{i,1}^r, x_{i,2}^r, \ldots, x_{i,|c|-r}^r$ seien die Originalpunkte von ξ_i^r. Um jeden von ihnen bestimmen wir ebenfalls ein Element $e_{i,j}^r$, so daß diese Elemente untereinander fremd sind, und daß $f(e_{i,j}^r) \subset \omega_i^r$ ist. In ω_i^r führen wir ein Polarkoordinatensystem ϱ, ψ mit ξ_i^r als Pol ein und ebenso in $e_{i,j}^r$ ein Polarkoordinatensystem R, φ mit $x_{i,j}^r$ als Pol. Wir betrachten nun für ein festes $e_{i,j}^r$ die Abbildung f; sie sei durch

$$\varrho = f_1(R, \varphi), \qquad \psi = f_2(R, \varphi)$$

gegeben, und sie habe in ξ_i^r den Grad a. Wir ersetzen sie durch die folgende Abbildung $\bar{f}$:

$$\begin{aligned}
&\bar{f}_1(R,\varphi) = f_1(R,\varphi), \quad \bar{f}_2(R,\varphi) = f_2(R,\varphi) \quad \text{für } R \geqq 2;\\
&\left.\begin{aligned}\bar{f}_1(R,\varphi) &= (R-1) f_1(R,\varphi) + (2-R)\\ \bar{f}_2(R,\varphi) &= (R-1) f_2(R,\varphi) + (2-R)\, a\varphi\end{aligned}\right\} \text{für } 2 \geqq R \geqq 1;\\
&\bar{f}_1(R,\varphi) = R, \quad \bar{f}_2(R,\varphi) = a\varphi \quad \text{für } 1 \geqq R.
\end{aligned}$$

Diese Gleichungen stellen in der Tat eine eindeutige Abbildung dar; um dies zu erkennen, hat man sich nur davon zu überzeugen, daß $\bar{f}_2(R, \varphi + 2\pi) - \bar{f}_2(R, \varphi)$ stets ein ganzes Vielfaches von 2π ist; für $R \geqq 2$ und $R \leqq 1$ ist das selbstverständlich, und für $2 > R > 1$ folgt es aus der Tatsache, daß $f_2(R, \varphi + 2\pi) - f_2(R, \varphi) = a \cdot 2\pi$ ist, weil die Abbildung $f(e_{i,j}^r)$ im Punkte ξ_i^r den Grad a hat, ξ_i^r also von dem Bilde jedes Kreises $R =$ konst. a-mal umlaufen wird.

Diese Abänderung von f nehmen wir in jedem $e_{i,j}^r$ vor und nennen die sich ergebende Abbildung $\bar{f}$; sie hat denselben Grad c wie f, da die Bedeckungen der außerhalb der ω_i^r gelegenen Punkte ungeändert geblieben sind. $\bar{f}$ hat die Eigenschaft, daß die Originalmengen der durch $\varrho < 1$ bestimmten offenen Teilmengen der ω_i^r die durch $R < 1$ bestimmten offenen Teile der $e_{i,1}^r, e_{i,2}^r, \ldots, e_{i,|c|-r}^r$ sind. Entfernen wir die genannten offenen Mengen aus den beiden Flächen, so wird aus der Originalfläche F_p eine Fläche F_p' vom Geschlecht p mit $\sum_{r=1}^{|c|-1} (|c| - r) v_r$ Rändern und aus der Bildfläche F_q eine Fläche F_q' vom Geschlecht q mit $\sum_{r=1}^{|c|-1} v_r$ Rändern, und F_p' ist durch $\bar{f}$ auf F_q' so abgebildet, daß die Randkurven in die Randkurven übergehen. Wir stellen nun von jeder der Flächen F_p', F_q' ein zweites Exemplar F_p'' bzw. F_q'' her, und bilden F_p'' auf F_q'' ebenfalls durch $\bar{f}$

ab. Fügen wir dann F'_p und F'_q längs entsprechender Ränder und ebenso F'_q und F''_q längs entsprechender Ränder zusammen, so entstehen zwei geschlossene, orientierbare Flächen F_P, F_Q, so daß F_P auf F_Q durch $\bar{f}$ eindeutig mit dem Grade c abgebildet ist. Dabei sind die Geschlechter dieser Flächen

$$P = 2p - 1 + \sum_r (|c| - r) v_r, \qquad Q = 2q - 1 + \sum_r v_r.$$

Ist $Q = 0$, so ist $q = 0$, $\sum_r v_r = 1$, also $\sum_r r v_r \leqq |c| - 1 < 2p - 2 + 2|c|$; ist $Q \geqq 1$, so ist nach dem oben genannten Kneserschen Satz

$$2p - 2 + |c| \sum_r v_r - \sum_r r v_r - 2q|c| + 2|c| - |c| \sum_r v_r \geqq 0,$$

d. h. es ist in jedem Fall

$$\sum_r r \cdot v_r \leqq (2p - 2) - |c|(2q - 2),$$

womit die Behauptung bewiesen ist.

Bemerkungen zu Satz XVIII. 1. Wenn die Bildmenge $f(F_p)$ als Riemannsche Fläche über F_q liegt, d. h. wenn die Umgebung jedes Punktes von F_q, der nicht Windungspunkt ist, genau c-mal glatt im positiven Sinne bedeckt wird, so gilt bekanntlich die „Hurwitzsche Formel“ [35])

$$\sum_r r w_r = (2p - 2) - c(2q - 2).$$

Unser Satz sagt also, daß die „Windungszahl“ $\sum_r r w_r$ einer solchen Abbildung den Höchstwert hat, der bei den vorliegenden p, q, c überhaupt möglich ist.

2. $w_{|c|-1}$ ist die Anzahl der Punkte auf F_q, die nur je einen Originalpunkt haben. Aus unserem Satz folgt

$$w_{|c|-1} \leqq \frac{2p - 2q|c|}{|c| - 1} + 2.$$

Satz XVI besagte:

$$w_{|c|-1} \leqq 2p + 2 - 2q.$$

Wenn nicht $|c| = 2$ und $q = 0$ ist, so ist die neue Schranke besser als die frühere. Ist $|c| = 2$, $q = 0$, so liefern beide Sätze

$$w_1 \leqq 2p + 2.$$

Diese Schranke läßt sich nicht verbessern; denn die Riemannsche Fläche der algebraischen Funktion

$$f(z) = \sqrt{z(z-1)(z-2)\ldots(z-2p)}$$

[35]) Kerékjártó, Vorlesungen über Topologie (Berlin 1923), S. 160.

hat das Geschlecht p, zwei Blätter und $2p+2$ Windungspunkte der Ordnung 1, nämlich $0, 1, 2, \ldots, 2p, \infty$.

3. Nach Satz XVa gibt es Abbildungen geschlossener Flächen mit $\sigma > j$; dabei ist in den dort angegebenen Beispielen $j > 4$. Es bleibt aber noch die Frage offen, ob es auch Abbildungen geschlossener Flächen mit $\sigma > j = 1$ gibt, — eine Frage, deren Beantwortung der Beweismethode des Satzes XVa, nämlich der Zurückführung auf Satz XVI, nicht zugänglich ist. Jetzt können wir diese Frage bejahen; es gilt nämlich

Satz XVb. *Es gibt eine Klasse von Abbildungen der geschlossenen orientierbaren Fläche F_2 (vom Geschlecht 2) auf die geschlossene orientierbare Fläche F_1 (vom Geschlecht 1) mit $\sigma > j = 1$.*

Beweis. $\sigma > 1$ ist gleichbedeutend mit $w_{|c|-1} = 0$ für alle Abbildungen der Klasse. Nach der soeben bewiesenen Formel (mit $p = 2$, $q = 1$)

$$w_{|c|-1} \leqq \frac{4-2|c|}{|c|-1} + 2$$

genügt daher die Angabe einer Abbildung f von F_2 auf F_1 mit $c = 4$ und $j = 1$. Eine solche Abbildung kann folgendermaßen hergestellt werden: Man zerlegt F_2 durch eine geeignete einfach geschlossene Kurve C in zwei Hälften F_2' und F_2'', von denen jede eine einmal berandete Fläche vom Geschlecht 1 ist. F_2' wird so auf F_1 abgebildet, daß der Rand in einen Punkt ξ übergeht und die Abbildung im übrigen eineindeutig ist; F_2'' wird so abgebildet, daß der Rand in denselben Punkt ξ übergeht und die Bildmenge $f(F_2'')$ im übrigen als 3-blätterige unverzweigte Überlagerungsfläche über F_1 liegt. $f(F_2)$ ist eine eindeutige Abbildung vom Grade 4. Sie hat den Index $j = 1$; denn jeder (nicht durch ξ gehende) geschlossene Weg auf F_1 ist Bild eines geschlossenen Weges auf F_2'.

(Eingegangen am 22. 6. 1929.)

On the Structure of Sets of Points of Classes One, Two, and Three.

Von

A. H. Blue in Iowa City (U.S.A.).

I. Introduction.

The Classification of Sets of Points.

The classification, defined by Lebesgue[1]), of sets of points is based upon the Baire classification of functions. Any set is closed or F of class α if it can be considered as the set $E\ (a \leqq f \leqq b)$ relative to a function f of class α, at most. A set is open or O of class α if it can be considered as the set $E\ (a < f < b)$ relative to a function f of class α, at most.

The following are some of the properties of the Lebesgue classification:

1. If a set is F of class α its complement is O of class α, and conversely;

2. A finite sum or product of sets F of class α, at most, gives a set F of class α, at most, and a finite sum or product of sets O of class α, at most, gives a set O of class α, at most;

3. The product of an enumerable infinity of sets F of class α, at most, is F of class α, at most, and the sum of an enumerable infinity of sets O of class α, at most, is O of class α, at most;

4. A set F of class α is O of class $\alpha + 1$, at most, and a set O of class α is F of class $\alpha + 1$, at most;

5. The sum of an enumerable infinity of sets F of class α, at most, is F of class $\alpha + 2$, at most, and the product of an enumerable infinity of sets O of class α, at most, is O of class $\alpha + 2$, at most.

---

[1]) Journal de Mathématiques 1905, pp. 156, 157.

For the sets which are both F and O of class α de la Vallée Poussin[2]) has introduced the term 'ambiguous' or in notation, A of class α. By the introduction of new and powerful methods de la Vallée Poussin was able to simplify and extend the theories of Baire and Lebesgue.

Two things are evident relative to the Lebesgue classification, a set belongs to two classes, open and closed, and its class depends directly on the class of a function. A classification which is not double and which depends only on the set itself is desirable. The sigma-delta-systems of Lebesgue[3]) and Hausdorff[4]) attain this latter aim. These classifications are based on the representation of sets in terms of open sets and closed sets.

In the Lebesgue classification the sets of class zero are open or O_0 and closed or F_0. The sets open of class one or O_1 are sums of sets F_0 and the sets closed of class one or F_1 are products of sets O_0. The sets open of class two or O_2 are sums of sets F_1 and the sets closed of class two or F_2 are products of sets O_1. This can be continued for all finite classes. For the first transfinite class ω the sets A'_ω and A''_ω are defined as the sum and product, respectively, of sets A of all finite classes. Then the sets open of class ω or O_ω are sums of sets A''_ω and the sets closed of class ω or F_ω are products of sets A'_ω. The process can then be continued for all transfinite classes.

The Hausdorff system is equivalent. Let G denote an open set and F a closed set. If E_δ denotes a product and E_σ a sum of sets E there are two systems of sets:

$$1.\quad G, G_\delta, G_{\delta\sigma}, G_{\delta\sigma\delta}, G_{\delta\sigma\delta\sigma}, \ldots,$$

$$2.\quad F, F_\sigma, F_{\sigma\delta}, F_{\sigma\delta\sigma}, F_{\sigma\delta\sigma\delta}, \ldots.$$

For the first system the classes are designated by the symbol (α) and for the second by $[\alpha]$. The open sets G are of class (1), the sets G_δ are of class (2), and so on for all finite classes (n). To the class (ω) belong the products of sets of classes $(1), (2), (3), (4), \ldots, (n), \ldots$. A sum of sets of class (ω) is of class $(\omega+1)$ and the process can be continued for all transfinite classes. In like manner the sets F are of class $[1]$, the sets F_σ are of class $[2]$, and the sets $F_{\sigma\delta}$ are of class $[3]$. The sets of class $[\omega]$ are sums of sets of classes $[1], [2], [3], \ldots, [n], \ldots$. These systems have been called delta- and sigma-systems, respectively. Certain relations between the two systems follow from the fact that an open set O is F_σ and a closed set F is G_δ.

---

[2]) Intégrales de Lebesgue, Paris 1916, pp. 150, 151.

[3]) de la Vallée Poussin, loc. cit. pp. 138, 139.

[4]) Math. Annalen 77, pp. 430—432.

These classifications of Lebesgue and Hausdorff are closely related. The sets of class $(\alpha+1)$ in the Hausdorff classification are open of class α in the Lebesgue system if α is even and are closed of class α if α is odd. The sets of class $[\alpha+1]$ in the Hausdorff classification are closed of class α if α is even and are open of class α if α is odd.

The classifications of Lebesgue and Hausdorff are successful in defining classes without relying on the Baire classification of functions. The double character of these classifications has already been pointed out, and in either of them the classes overlap. From the properties cited for the Lebesgue classification it follows that a set open of class α is closed of class $\alpha-1$, α, or $\alpha+1$ and a set closed of class α is open of class $\alpha-1$, α, or $\alpha+1$. These sets, though they have the same open class, or the same closed class, are distinctly different. To make a distinction between these sets is a part of the purpose of this paper.

It is only natural to suppose that there are certain structural characteristics peculiar to the sets of a particular class. What are those structural characteristics possessed in common by all the sets of a given class? In this investigation an answer to this question has been found for sets of classes one and two, and in part for sets of class three.

II. The Structure of Sets of Points.

The Type of a Set.

In the Lebesgue classification the class of a set is defined in terms of its representation with closed sets and open sets. If the class be defined in terms of the simplest representation with closed sets and open sets, that is, the representation requiring not more alternate sums and products than any other such representation, then the classes are distinct and do not overlap.

The distinction to be made between those sets which are of the same open class but are of different closed classes is based on the notion of the "type" of a set. If a set is open of class α and is closed of class β it is of "type" (α, β), and conversely.

Of two sets A and B of types (α', β') and (α, β), respectively, A is of lower type if $\alpha' \leqq \alpha$ and $\beta' \leqq \beta$ and the equalities do not hold simultaneously. Of two sets A and B of types (α, β) and (β, α), respectively, it cannot be said that either is of lower type than the other. However, the two types are distinct.

Corresponding to the properties cited for the Lebesgue classification are the following properties of types:

1. Every finite sum or product of sets of type (α', β'), where $\alpha' \leqq \alpha$ and $\beta' \leqq \beta$, is of type (α, β), at most;

2. Every enumerably infinite sum or product of sets of type (α', β'), where $\alpha' \leqq \alpha$ and $\beta' \leqq \beta$, is of type $(\alpha + 1, \beta + 1)$, at most;

3. The removal of a subset of type (α, β) from a set of the same type leaves a subset of type (γ, γ), at most, where γ is the larger of α and β;

4. The removal of a subset of type (α, β) from a set of lower type leaves a subset of type (β, α), at most.

A Property of Sets of Form G_δ.

For two sets which are of form G_δ the following theorem holds:

Theorem 1. *If two sets of form G_δ are everywhere dense*[5]) *in a perfect set they have a common point.*

Assume the contrary, that two sets A and B of form $\prod G_m$ and $\prod \overline{G}_m$, respectively, have no common points and are everywhere dense in a perfect set H. Since A is a subset of G_m and B is a subset of $\overline{G}_m$, both G_m and $\overline{G}_m$ must be everywhere dense in H while their complementary closed sets are nowhere dense in H. If $U_m = G_m - \overline{G}_m$ and $V_m = \overline{G}_m - G_m$, these sets are nowhere dense in H. By definition, A is the product of sets $(U_m + G_m \cdot \overline{G}_m)$. By hypothesis, $A \cdot B = \prod G_m \cdot \overline{G}_m$ is a null set. Hence, any point of A must be in every U_m for m sufficiently large. Therefore, A may be expressed as the sum of sets $A_n = \prod\limits_{m=n}^{\infty} U_m$ each nowhere dense in H. Therefore, A is of the first category relative to H. The complement of A is evidently of the first category relative to H. Since a set and its complement cannot both be of the first category, the desired contradiction is obtained.

From this theorem it is easy to demonstrate the following propositions:

1. If a subset of any set A of form $F_{\sigma\delta}$ is of the second category relative to a perfect set H in which a subset of any set B of form G_δ is everywhere dense, then A and B have a common point;

2. There exists a perfect set H relative to which the subset $A \cdot H$ of any set A of type $(2, 1)$ or $(2, 2)$ is of the second category.

By definition, a set of type $(2, 1)$ is of form $F_{\sigma\delta}$ and G_δ and a set of type $(2, 2)$ is of form $F_{\sigma\delta}$ and $G_{\delta\sigma}$. Hence, the complement, B, is of form $\overline{G}_{\delta\sigma}$.

---

[5]) Acta Mathematica **30** (1906), pp. 10—12.

Assume the contrary. Then $B \cdot H$ is everywhere of the second category relative to H, and by Theorem 1 any subset G_δ of A is nowhere dense in any perfect set H. But such a set is enumerable. Therefore, A is enumerable and of form F_σ, contrary to the hypothesis.

Structure of the Sets of Type (1, 1).

Theorem 2. *The necessary and sufficient condition that any set A which is neither open nor closed be of type* $(1, 1)$ *is that if a subset of A or its complement B be everywhere dense in a perfect set H then $B \cdot H$ or $A \cdot H$, respectively, is nowhere dense in H.*

By definition, both the set A and its complement B of type $(1, 1)$ are of form G_δ. Any perfect set H is of form G_δ and the subsets $A \cdot H$ and $B \cdot H$ are then of the same form. By Theorem 1, the condition is necessary, since A and B have no common point.

Consider the closed derived set of order zero, A^0, which is the sum of a perfect set H and a reducible set[6]) N. The set $A \cdot H$ is everywhere dense in H and by hypothesis $B \cdot H$ is nowhere dense in H. If G be the open set determined by all those portions of H containing no point of B, the set $G \cdot H$ is a subset of A. If $A_1 = A - G \cdot H$, the set $A_1 \cdot H_1$ is everywhere dense in the perfect subset H_1 of A_1^0. If G_1 be the open set determined by all those portions of H_1 containing no point of $B \cdot H_1$, $G_1 \cdot H_1$ is a subset of A_1 and of A. Let $A_2 = A_1 - G_1 \cdot H_1$.

The continuation of this process gives rise to a sequence of closed sets A_α^0 such that each is a subset of its predecessors. Let A_μ^0 be the closed set of points common to all the sets A_α^0, where $\alpha < \mu$ of the second kind. Defining the corresponding sets H_μ and G_μ the process may be continued for transfinite ordinals.

By the fundamental principle of Cantor[7]) there exists a least number β such that the sets $A_\beta^0, A_{\beta+1}^0, A_{\beta+2}^0, \ldots$, are identical. It follows that the corresponding perfect sets $H_\beta, H_{\beta+1}, H_{\beta+2}, \ldots$, are null sets, since $G_\beta \cdot H_\beta$ must be a null set when $A_\beta = A_{\beta+1}$ and G_β is not a null set if H_β exists. Also A_β is a reducible set N_β when H_β is a null set. By definition, the sets $G \cdot H, G_1 \cdot H_1, G_2 \cdot H_2, \ldots$, are non-overlapping. Every open set is a sum of closed sets and every set $G_\alpha \cdot H_\alpha$ is of form F_σ. The reducible set N_β is enumerable and also of form F_σ. The set A is the sum of the sets $G_\alpha \cdot H_\alpha$ and the reducible set N_β. Therefore, A is of form F_σ. By the same argument the complement B is also of form F_σ. Therefore, A_σ is of

[6]) Annali di Matematica (3) **3**, p. 37.

[7]) Math. Annalen **17** (1880), S. 357.

form F_σ and G_δ and of type $(1,1)$. by definition. Therefore, the condition is sufficient.

From this theorem it is evident that a reducible set, which is nowhere dense in every perfect set is not of higher type than $(1,1)$. That is, a reducible set must be of type $(1,0)$ or $(1,1)$.

The Decomposition of Sets of Type (1,1).

Theorem 2 furnishes a procedure for the decomposition of any set of type $(1,1)$ into a sum of disjoined sets of form $F \cdot G$. Let $A_0 = A$ be any set of type $(1,1)$ with its complement B. The set $A_0 \cdot H_0$ is every where dense in the perfect subset H_0 of A_0^0, and $B \cdot H_0$ is nowhere dense in H_0. In every portion of H_0 there is a portion containing no point of $B \cdot H_0$. By definition, G_0 is the open set determined by all those portions of H_0 containing no point of $B \cdot H_0$ and the set $G_0 \cdot H_0$ is a subset of A_0. Let $A_1 = A - G_0 \cdot H_0$ and repeat the same process on A_1. The set $A_1 \cdot H_1$ is everywhere dense in the perfect subset H_1 of A_1^0, and $B \cdot H_1$ is nowhere dense in H_1. Let G_1 be the open set determined by all those portions of H_1 containing no point of $B \cdot H_1$. Then the set $G_1 \cdot H_1$ is a subset of A_1 and of A, and the sets $G_0 \cdot H_0$ and $G_1 \cdot H_1$ are disjoined. Let $A_2 = A_1 - G_1 \cdot H_1$ and continue the decomposition with A_2. As was shown in Theorem 2 this decomposition must end, since it gives rise to a decreasing sequence of closed sets $A_0^0, A_1^0, A_2^0, \ldots$. The set A_0 is thus decomposed into the sum of disjoined sets $G_\alpha \cdot H_\alpha$ and a reducible set N_β, where $\alpha < \beta < \Omega$. Therefore.

$$A_0 = A = G_0 \cdot H_0 + G_1 \cdot H_1 + G_2 \cdot H_2 + \ldots + G_\alpha \cdot H_\alpha + \ldots + N_\beta .$$

The Structure of Sets of Type (1,2) and (2,1).

It is evident from the condition of Theorem 2 that sets of higher type than $(1,1)$ must all possess the property of being everywhere dense with their complements in some perfect set.

The following proposition characterizes the sets of type $(1,2)$:

The necessary and sufficient condition that any set A of form F_σ, with the complementary set B, be of type $(1,2)$ is that there exist a perfect set H in which both the sets $A \cdot H$ and $B \cdot H$ are everywhere dense.

If no such perfect set exists the set is of type $(1,1)$, at most. Therefore, the condition is necessary.

If there exists such a perfect set, A cannot be open or closed, or of type $(1,1)$, by Theorem 2. Any set of form F_σ is of type $(1,2)$, at most. Therefore, the condition is sufficient.

Since the complement of a set of type $(1,2)$ is a set of type $(2,1)$, there is the analogous proposition for sets of type $(2,1)$:

The necessary and sufficient condition that any set A of form G_δ, with the complementary set B, be of type $(2,1)$ is that there exist a perfect set H in which both the sets $A \cdot H$ and $B \cdot H$ are everywhere dense.

The characterization of sets of type $(1,2)$ at the same time exposes the character of their complementary sets of type $(2,1)$.

The Decomposition of Sets of Type (1, 2).

Let A_0 be any set of type $(1,2)$ with its complement B. If B is everywhere dense in the perfect subset H_0 of A_0^0, then A_0 cannot fill any portion of H_0. Hence, A_0 is the sum of closed sets nowhere dense in H_0 and a reducible set. If B is not everywhere dense in H_0, there is at least one portion of H_0 containing no point of B. As in the case of a set of type $(1,1)$ a subset $G_0 \cdot H_0$ may be removed from A_0 leaving a subset A_1. The decomposition may then be repeated with A_1. It is evident that B must be everywhere dense in some perfect set H_α. Then A_0 is the sum of closed sets nowhere dense in H_α, and the sets of points common to the open sets G_α and the perfect sets H_α and a reducible set. The open sets will be null sets if B is everywhere dense in H_0.

For the decomposition of sets of type $(1,2)$ the following proposition may then be stated:

Any set of type $(1,2)$ is the sum of closed sets nowhere dense in a perfect set, a reducible set, and the sum of non-overlapping sets $G_\alpha \cdot H_\alpha$, where $\alpha < \beta < \Omega$.

The Structure of Sets of Form $F_{\sigma\delta}$.

Theorem 3. *If any set A of form $F_{\sigma\delta}$ has a subset $A \cdot H$ of the second category relative to a perfect set H, then $A \cdot H$ is the sum of sets $A_1 \cdot H$ and G_δ of the first and second categories, respectively, relative to H.*

The set $A \cdot H$ is the common subset of sets F_σ which must also be of the second category relative to H and everywhere of the second category relative to at least one portion of H. For each set F_σ, in a portion H_2 of H_1 at least one of the closed sets F is dense and fills a portion of H_2. It follows that each set F_σ is a sum of closed sets nowhere dense in H_1 and an open set G which is the sum of all those open sets determined by the portions filled by the closed sets F. These open sets are everywhere dense in H_1. Hence, $A \cdot H$ is the sum of the subset $A_1 \cdot H$ of the first category relative to H and the subset G_δ everywhere of the second category relative to H_1.

Let A be any set of form $F_{\sigma\delta}$. If A is of the second category relative to the fundamental set P it is the sum of a subset A_1 of the first category relative to P and a set G_δ, by Theorem 3.

If one of the sets F_σ of which A is the common subset is of the first category relative to P then A is also and the procedure is the same as for the subset A_1 in the following argument.

Suppose A_1 is not of form F or F_σ, it is then of form $F_{\sigma\delta}$. Every closed set is the sum of a perfect set H and a reducible set. Hence, taking the product of such sets, $A_1 = \Pi(\Sigma H_{n_1} + N) = \Pi \Sigma H_{n_1} + N_1$. Except for the enumerable subset N_1, A_1 is a subset of the sum of non-overlapping perfect sets H_{n_1}, of diameter less than a given positive number ε_1, each nowhere dense in P. Consider the category of $A_1 \cdot H_{n_1}$ relative to the perfect set H_{n_1}. By Theorem 3, we may remove a set G_1 from A_1 if it is of the second category relative to H_{n_1}. Therefore, A_1 is the sum of an enumerable set N_1, a subset A_2 of the first category relative to each set H_{n_1} and a set of form $G_{1\delta}$ or $G_{1\delta\sigma}$, respectively, when A_1 is of the second category relative to a finite number or an enumerable infinity of the perfect sets H_{n_1}.

Suppose A_2 is not of form F or F_σ, it is then of form $F_{\sigma\delta}$. Except for an enumerable subset N_2, A_2 is a subset of the sum of non-overlapping perfect sets $H_{n_1 n_2}$ each nowhere dense in H_{n_1} and of diameter less than a given positive number ε_2. Consider the category of $A_2 \cdot H_{n_1 n_2}$ relative to each perfect set $H_{n_1 n_2}$. As before, A_2 is the sum of an enumerable set N_2, a subset A_3 of the first category relative to each set $H_{n_1 n_2}$ and a set of form $G_{2\delta}$ or $G_{2\delta\sigma}$, respectively, when A_2 is of the second category relative to a finite or an enumerable infinity of sets $H_{n_1 n_2}$.

Suppose there has been chosen a monotonic decreasing sequence of positive numbers ε_r with the limit zero. The continuation of the foregoing process gives rise to sequences of perfect sets such as the following:

$$H_{n_1},\ H_{n_1 n_2},\ H_{n_1 n_2 n_3},\ \ldots,\ H_{n_1 n_2 n_3 \ldots n_r},\ \ldots,$$

where each set is a subset of its predecessor and is of diameter less than the corresponding ε_r. Each of these sequences must end with a null set, when one of the sets contains no point and the corresponding closed set is finite or enumerable, or the sequence must determine a point of a residual set A_ω.

Consider the sequence of subsets of $A, A_1, A_2, A_3, \ldots, A_\alpha, \ldots$, where $\alpha < \omega + 1$. Suppose any set A_α of this sequence is of form F, F_σ or G_δ. Then the process of the decomposition of A should stop with that set A_α. The set A is then the sum of an enumerable set of form F or F_σ, a set of form $G_{\delta\sigma}$, that is, the sum of sets G_δ removed at each step, and a

set A_α of form F, F_σ or G_δ. The enumerable set is open of class one, the set of form $G_{\delta\sigma}$ is open of class two, at most, and the set A_α is open of class two, at most. Therefore, A is open of class two, at most. By hypothesis, A is of form $F_{\sigma\delta}$ and is, therefore, closed of class two, at most. Therefore, A is of type $(2, 2)$, at most.

The Structure of the Residual Set A_ω.

Suppose there is no set A_α of the sequence of form F, F_σ, or G_δ. Then the residual set A_ω exists and is of form $F_{\sigma\delta}$.

Theorem 4. *If the residual set A_ω of form $F_{\sigma\delta}$ is dense in the fundamental set P it is of type* $(3, 2)$.

Assume the contrary, that the complementary set B is also of form $F_{\sigma\delta}$ or $\prod\limits_m B_m$ where $B_m = \sum\limits_n F_{mn}$ and B_{m+1} is a subset of B_m. Now A is of the first category relative to P. It follows that B, and consequently each set B_m, is everywhere of the second category relative to P. Hence, for each set B_m, in a portion of P at least one of the closed sets F_{mn} is dense and fills a portion P_m of P.

By hypothesis, A is dense in P, that is, A is everywhere dense in a portion K of P. The set B_1 fills a portion P_1 of K and has a perfect subset $H_{n_1 n_2 \ldots n_{r_1}}$. For this it is sufficient to choose r_1 so that the diameter ε_{r_1} of $H_{n_1 n_2 \ldots n_{r_1}}$ is less than one half the diameter of P_1. In the portion of P_1 of diameter ε_{r_1} determined by $H_{n_1 n_2 \ldots n_{r_1}}$, the set B_2 fills a portion of P_2. There exists a number r_2, such that the perfect set $H_{n_1 n_2 \ldots n_{r_1} \ldots n_{r_2}}$ is a subset of B_2. In general, for a set B_m there exists a number r_m such that the perfect set $H_{n_1 n_2 \ldots n_{r_1} \ldots n_{r_2} \ldots n_{r_m}}$ is a subset of B_m and of every $B_{m'}$ preceding B_m.

But the sequence of perfect sets

$$H_{n_1 n_2 \ldots n_{r_1} \ldots n_{r_2} \ldots n_{r_m}} \qquad (m = 1, 2, \ldots)$$

determines a point of A_ω. This point is a point of B since every set B_m contains a perfect set of this sequence. But B is the complement of A_ω. Therefore, the assumption that B is also of the form $F_{\sigma\delta}$ is false, and A_ω cannot be of type $(2, 2)$.

Since any set of the form $F_{\sigma\delta}$ is (by construction) of the type $(3, 2)$ at most, the theorem follows.

University of Iowa, January 19, 1929.

(Eingegangen am 12. 2. 1929.)

Über die Lösungen der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus.

Von

Willy Feller in Kiel.

Bekanntlich kann die allgemeine lineare partielle Differentialgleichung zweiter Ordnung mit *zwei unabhängigen Veränderlichen* auf eine Normalform gebracht werden, welche für alle Untersuchungen besonders geeignet ist, da sie in den zweiten Ableitungen mit der Laplaceschen Gleichung übereinstimmt. Für diese Gleichung hat Herr Lichtenstein in verschiedenen Arbeiten[1]) die Randwert- bzw. die Eigenwertaufgabe in größter Allgemeinheit gelöst, und ihre Lösungen verhalten sich in mancher Hinsicht ähnlich wie die Potentialfunktionen. Die dabei benutzten Methoden können aber nicht auf den Fall von mehreren unabhängigen Veränderlichen übertragen werden; indessen hat Herr Sternberg[2]) die Randwertaufgabe bei drei Veränderlichen auf eine lineare Integralgleichung zurückgeführt, indem er von Integralausdrücken ausgeht, die ganz den bekannten Raum- und Flächenpotentialen für die Laplacesche Gleichung nachgebildet sind.

In dieser Arbeit soll gezeigt werden, wie sich fast alle klassischen Sätze über die Eigenschaften der Potentialfunktionen ohne weiteres auf die Lösungen der allgemeinen selbstadjungierten Differentialgleichung zweiter Ordnung vom elliptischen Typus übertragen, wenn man zu ihrer Formulierung die naturgemäß mit der Gleichung verbundene *Riemannsche Maßbestimmung* benutzt. Es bleiben dann fast alle Formeln mit geringfügigen Änderungen bestehen, so ein Analogon zum *Mittelwertsatz*, die Integraldarstellungen und Sprungrelationen, die *Gaußsche geometrische Deutung* für die Doppel-

[1]) Nähere Literaturangaben in der Encyklopädie der math. Wissensch. II, C, 12: L. Lichtenstein, Neuere Entwicklung der Theorie partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus (1924).

[2]) W. Sternberg, Über die lineare elliptische Differentialgleichung zweiter Ordnung mit drei unabhängigen Veränderlichen. Math. Zeitschr. 21 (1924), S. 286—311.

belegung usw. Viele von diesen Eigenschaften übertragen sich auch auf die *nicht selbstadjungierten* Differentialgleichungen. So ergibt sich z. B. ein äußerst einfacher Beweis der sog. *Harnackschen Konvergenzsätze* für *beliebige* Gleichungen.

Alle folgenden Entwicklungen werden naturgemäß nach Möglichkeit invariant geschrieben gegenüber allen Koordinatentransformationen, und der selbstadjungierte Differentialausdruck wird daher als zweiter Differentialparameter der betreffenden Maßbestimmung aufgefaßt. Vorkenntnisse über Riemannsche Geometrie werden übrigens nicht vorausgesetzt, bloß im folgenden Abschnitt muß das Volumelement einer im n-dimensionalen Raume eingebetteten $(n-1)$-dimensionalen Mannigfaltigkeit entwickelt werden. Es wird gleich allgemein der Fall von n unabhängigen Veränderlichen ins Auge gefaßt, da sich die Darstellung nicht umständlicher gestaltet als für $n = 3$ und der erste Teil doch in dieser Allgemeinheit notwendig ist.

§ 1.

Das Volumelement eingebetteter Mannigfaltigkeiten im Riemannschen Raume.

Bekanntlich definiert man das Volumelement im Raume mit dem (definiten) Linienelement

$$ds^2 = \sum_{i,k=1}^{n} g_{ik}\, dx_i\, dx_k \qquad (g_{ik} = g_{ki}) \tag{1}$$

durch

$$dV = \sqrt{g}\, dx_1 \ldots dx_n,$$

wobei mit g die wesentlich positive Determinante der Matrix (g_{ik}) bezeichnet wurde. Eine in diesem Raume eingebettete $(n-1)$-dimensionale Mannigfaltigkeit

$$F(x_1, \ldots, x_n) = \sigma = \text{konst.} \tag{2}$$

können wir auch durch $n-1$ unabhängige Parameter $\lambda_1, \ldots, \lambda_{n-1}$ in der Form

$$x_i = x_i(\lambda_1, \ldots, \lambda_{n-1}) \qquad (i = 1, \ldots, n) \tag{2'}$$

bestimmen. Die Maßbestimmung auf dieser Hyperfläche erhalten wir, indem wir diese Größen in das Linienelement (1) einsetzen; so erhalten wir zugleich auch das Volumelement dieser Mannigfaltigkeit. Für das Folgende müssen wir jedoch, ähnlich wie im euklidischen Falle, dieses Volumelement direkt durch die Koeffizienten g_{ik} und die Größen (2') ausdrücken.

Es möge wie üblich g^{ik} das durch die Determinante g dividierte algebraische Komplement von g_{ik} in der Matrix (g_{ik}) bezeichnen; ferner

bezeichnen wir mit D_i $(i = 1, \ldots, n)$ die n Funktionaldeterminanten der Koordinaten $x_1, \ldots, x_n$ nach den Parametern $\lambda_1, \ldots, \lambda_{n-1}$, also

$$D_i = (-1)^{i-1} \frac{\partial(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)}{\partial(\lambda_1, \ldots, \lambda_{n-1})}.$$

Wir zeigen, daß das gesuchte *Volumelement der Mannigfaltigkeit* (2′) geliefert wird durch den Ausdruck[3])

$$(3) \qquad do = \sqrt{g} \sqrt{\sum_{l,k=1}^{n} g^{lk} D_l D_k} \cdot d\lambda_1 \ldots d\lambda_{n-1}.$$

Zum Beweise zeigen wir zunächst, daß dieser Ausdruck unabhängig ist sowohl von der besonderen Wahl der Parameter $\lambda_1, \ldots, \lambda_{n-1}$, als auch vom speziellen Koordinatensystem $x_1, \ldots, x_n$ des Raumes. Sodann genügt offenbar der Nachweis, daß bei einer *besonderen* Wahl des Koordinatensystems und der Parameter das gesuchte Volumelement in der Tat durch (3) dargestellt wird.

Die Größen g^{lk} und g sind nur vom ursprünglichen Raume, nicht von der Hyperfläche abhängig, während sich die Funktionaldeterminanten D_i bei einer Transformation der Parameter λ_i mit der Determinante der Transformation multiplizieren: daher bleibt die Größe (3) invariant gegenüber allen Transformationen der Parameter.

Es bleibt zu zeigen, daß der Ausdruck

$$\sqrt{g} \sqrt{\sum_{l,k=1}^{n} g^{lk} D_l D_k}$$

unabhängig ist von der Wahl des Koordinatensystems x_i im Raume. Der Kürze halber führen wir die Bezeichnung

$$F_i = \frac{\partial F}{\partial x_i} \qquad (i = 1, \ldots, n)$$

ein. Es ist, wie man leicht bestätigen kann,

$$F_i : F_k = D_i : D_k$$

und daher

$$(4) \qquad \frac{F_i}{\sqrt{\sum_{l,k=1}^{n} g^{lk} F_l F_k}} = \frac{D_i}{\sqrt{\sum_{l,k=1}^{n} g^{lk} D_l D_k}}$$

Der Nenner auf der linken Seite ist, wie leicht nachzurechnen, eine Invariante (nämlich der sog. erste Differentialparameter von Beltrami der

[3]) Diese Behauptung reduziert sich bekanntlich im Falle einer euklidischen Maßbestimmung auf die *Identität von Lagrange* und stellt daher eine Verallgemeinerung derselben dar.

Funktion F); um das Verhalten des Ausdrucks $\sqrt{\sum g^{ik} D_i D_k}$ bei Koordinatentransformationen zu erkennen, fassen wir für den Augenblick auch σ als unabhängige Veränderliche auf, so daß uns die n-Gleichungen (2) und (2′) eine Transformation der Veränderlichen $x_1, \ldots, x_n$ in neue Veränderliche $\sigma, \lambda_1, \ldots, \lambda_{n-1}$ liefern. Die Funktionaldeterminante dieser Transformation berechnet sich mit Hilfe von (4) zu

$$D = \frac{\partial(x_1, \ldots, x_n)}{\partial(\sigma, \lambda_1, \ldots, \lambda_{n-1})} = \sum_{i=1}^{n} \frac{\partial x_i}{\partial \sigma} D_i = \sqrt{\frac{\sum g^{ik} D_i D_k}{\sum g^{ik} F_i F_k}} \sum_{i=1}^{n} \frac{\partial x_i}{\partial \sigma} F_i = \sqrt{\frac{\sum g^{ik} D_i D_k}{\sum g^{ik} F_i F_k}}.$$

Es ist mithin

$$\sqrt{\sum g^{ik} D_i D_k} = D \sqrt{\sum g^{ik} F_i F_k},$$

d. h. die linke Seite verhält sich bei Transformationen der x — da ja die Wurzel rechts eine Invariante ist — wie die Determinante D. Daher verhält sich in der Tat die Größe

$$\sqrt{g} \sqrt{\sum g^{ik} D_i D_k}$$

invariant gegenüber allen Koordinatentransformationen. Damit sind beide Invarianzeigenschaften des Ausdrucks (3) bewiesen, und es ist leicht zu zeigen, daß er das Volumelement der Mannigfaltigkeit $F = \sigma$ darstellt. In einem hinreichend kleinen Gebiete können wir nämlich das Koordinatensystem so wählen, daß diese Hyperfläche etwa durch die Gleichung $x_n =$ konst. dargestellt wird. Fassen wir dann die übrigen Koordinaten $x_\alpha = \lambda_\alpha$, $\alpha = 1, \ldots, n-1$, als Parameter auf, so wird der Definition nach das Volumelement gleich $\sqrt{\gamma}\, dx_1 \ldots dx_{n-1}$, wenn γ die Determinante der $(n-1)$-reihigen Matrix $(g_{\alpha\beta})$ bezeichnet $(\alpha, \beta = 1, \ldots, n-1)$. Es ist also $\gamma = g g^{nn}$. Andererseits wird offenbar $D_\alpha = 0$ für $\alpha = 1, \ldots, n-1$ und $D_n = 1$, so daß auch die rechte Seite von (3) $do = \sqrt{g g^{nn}}\, dx_1 \ldots dx_{n-1}$ ergibt. Da aber dieser Ausdruck invariant ist gegenüber allen Transformationen der Koordinaten und der Parameter, so stellt er tatsächlich das Oberflächenelement dar.

§ 2.

Die Greenschen Formeln und die Grundlösung.

Um unser Ziel zu erreichen, müssen wir auch den *Greenschen Formeln* eine invariante Fassung geben, ähnlich wie sie in der Flächentheorie üblich ist.

Wir betrachten zunächst einen *selbstadjungierten* elliptischen Differentialausdruck mit mehr als zwei unabhängigen Veränderlichen. Wir werden ihn

im folgenden stets in der Form des zum Linienelement (1) gehörigen zweiten *Beltramischen Differentialparameter*

$$(5) \qquad \Delta_2 u = \frac{1}{\sqrt{g}} \sum_{i,k=1}^{n} \frac{\partial}{\partial x_k} (g^{ik} \sqrt{g}\, u_i), \qquad g = |g_{ik}| = \frac{1}{|g^{ik}|}, \qquad u_i = \frac{\partial u}{\partial x_i}$$

gegeben denken. Man kann im Falle von $n > 2$ unabhängigen Veränderlichen jeden selbstadjungierten Differentialausdruck

$$\sum_{i,k=1}^{n} \frac{\partial}{\partial x_k} (A^{ik} u_i)$$

in dieser Form schreiben, indem man $g^{ik} = A^{\frac{1}{2-n}} A^{ik}$ setzt. $\left(A = \frac{1}{|A^{ik}|}\right)$ und den Ausdruck mit $\frac{1}{\sqrt{g}}$ multipliziert. Die Koeffizienten des Linienelements (1) sind dann die algebraischen Komplemente von (g^{ik}) dividiert durch die Determinante. Im übrigen ist diese Schreibweise für das Folgende nicht wesentlich; läßt man nämlich den Faktor $\sqrt{g}$ fort, so kommt in den zu entwickelnden Formeln bloß überall eine Funktionaldeterminante als Faktor vor. Die Koeffizienten g^{ik} mögen im ganzen betrachteten Bereiche T mit hinreichenden Stetigkeitseigenschaften versehen sein, z. B. stetige, einer Hölderschen Bedingung genügende zweite Ableitungen besitzen (auf eine möglichst weitgehende Reduktion der Voraussetzungen kommt es uns im folgenden nicht an); die Diskriminante g soll in diesem Bereiche wesentlich positiv sein.

Unseren Differentialausdruck multiplizieren wir nun mit dem Volumelement $dV = \sqrt{g}\, dx_1 \ldots dx_n$ und integrieren über ein (beschränktes) Gebiet G; dieses soll ganz in T liegen und durch eine zweimal stetig differenzierbare Hyperfläche F begrenzt sein, die wir uns durch die Gleichungen (2′) gegeben denken. Durch partielle Integration erhalten wir in bekannter Art die sog. *zweite Formel von Green*, die wir in folgender Form schreiben können:

$$(6) \qquad \iiint_G (v\, \Delta_2 u - u\, \Delta_2 v)\, dV = \iint_F \left(v \frac{\partial u}{\partial n} - u \frac{\partial v}{\partial n}\right) do.$$

Dabei wurde, wie auch im folgenden, der Übersichtlichkeit halber durch ein dreifaches Integrationszeichen ein n-dimensionales Gebietsintegral bezeichnet, durch das Doppelintegral das $(n-1)$-fache Integral über dessen Rand. Ferner ist das Oberflächenelement do durch (3) definiert, während

$$(7) \qquad \frac{\partial}{\partial n} = \sum_{l,k=1}^{n} g^{lk} \frac{D_k}{\sqrt{\Sigma g^{ij} D_i D_j}} \frac{\partial}{\partial x_l}$$

ist[4]); es ist das die ***normierte Differentiation in der Richtung der Normale auf die Randfläche*** im Sinne unserer Riemannschen Maßbestimmung, und zwar ist dadurch bei geschlossenen Hyperflächen die ***äußere*** Normale festgelegt. Man kann das leicht bestätigen, wenn man die Beziehung

$$\sum_{k,s} D_k \frac{\partial x_k}{\partial \lambda_s} d\lambda_s = 0$$

beachtet, die sich aus den Gleichungen (4) und (2) ergibt. Die allgemeine Greensche Formel hat demnach so geschrieben genau dieselbe Form wie diejenige für den Laplaceschen Ausdruck. Insbesondere ist die so normierte Differentiation in der Normalenrichtung bei einer Kugel mit dem Radius r gleich der Differentiation nach dem Radius r, wie man direkt aus den Differentialgleichungen der geodätischen Linien bestätigt, oder indem man die Entfernung r als unabhängige Veränderliche in das Linienelement einführt.

Den allgemeinsten ***nicht selbstadjungierten*** Differentialausdruck schreiben wir in der Form

$$(8) \qquad L(u) = \Delta_2 u + \sum_{i=1}^{n} b_i u_i + cu = \frac{1}{\sqrt{g}} \sum_{i,k} \frac{\partial}{\partial x_k} (g^{ik} \sqrt{g}\, u_i) + \sum_i b_i u_i + cu;$$

als den dazu ***adjungierten*** Ausdruck müssen wir in unserer Schweibweise

$$(9) \qquad M(v) = \Delta_2 v - \frac{1}{\sqrt{g}} \sum_i \frac{\partial}{\partial x_i} (b_i \sqrt{g}\, v) + cv$$

ansehen. Wie vorhin erhalten wir durch Produktintegration

$$(10) \qquad \iiint_G \{v L(u) - u M(v)\}\, dV = \iint_F \left\{ v \frac{\partial u}{\partial n} - u \frac{\partial v}{\partial n} + N u v \right\} do$$

mit

$$(11) \qquad N = \sum_i \frac{b_i D_i}{\sqrt{\sum_{l,k} g^{lk} D_l D_k}} = \sum_i b_i \cos(n, x_i);$$

$\frac{\partial}{\partial n}$ bezeichnet wiederum die Differentiation in der (äußeren) Normalenrichtung der Randfläche.

Man kann nun diese Formeln mit genau denselben Ergebnissen anwenden, wie man es in der Potentialtheorie tat, wenn man eine ***Grundlösung*** der Differentialgleichung benutzt. Eine solche hat bekanntlich

---

[4]) Diese Definition unterscheidet sich von der üblichen Differentiation in der „*Transversalenrichtung*" durch den Faktor $\frac{1}{\sqrt{g}\sqrt{\sum g^{ij} D_i D_j}}$; vgl. z. B. J. Hadamard, Lectures on Cauchys Problem in linear partial differential Equations, New Haven 1923.

E. E. Levi[5]) mit Hilfe der Fredholmschen Theorie konstruiert; ihr *Hauptteil* ist

$$\frac{1}{\left\{\sum_{i,k} g_{ik}(x_i-\xi_i)(x_i-\xi_k)\right\}^{\frac{n-2}{2}}}.$$

Es erspart uns manche Abschätzungen, wenn wir statt dessen eine Grundlösung benutzten, deren Hauptteil

$$\frac{1}{s^{n-2}}$$

ist, wobei s die geodätische Entfernung des Punktes (x) vom Aufpunkt (ξ) ist. Die Konstruktion ist natürlich dieselbe, sogar etwas einfacher, als bei E. E. Levy; wesentlich ist bloß zu zeigen, daß $\Delta_2\left(\frac{1}{s^{n-2}}\right)$ von der Größenordnung $\frac{1}{s^{n-1}}$ ist[6]).

Diese Grundlösung $\Gamma(x,\xi)$ der selbsadjungierten Gleichung ist außerhalb des Punktes (ξ) zweimal stetig nach den Koordinaten x_i differenzierbar; wir können sie, solange wir uns auf einen begrenzten Bereich be-

---

[5]) E. E. Levi, Sulle equazioni lineari totalmente ellittiche alle derivate parziali. Rend. del circ. mat. Palermo **24** (1907), S. 275—317, insb. S. 311 ff.

[6]) Man sieht das am einfachsten, indem man die *kanonischen Koordinaten* von Riemann benutzt. Dieses Koordinatensystem $(\mathfrak{x}_1, \ldots, \mathfrak{x}_n)$ ist so gewählt, daß die Gleichungen der vom Punkte $(\xi_1, \ldots, \xi_n)$ ausgehenden geodätischen Linien mit der Bogenlänge s als Parameter die Form

$$\mathfrak{x}_\alpha = c_\alpha s, \qquad c_\alpha = \text{konst.} = \left(\frac{d\mathfrak{x}_\alpha}{ds}\right)_{s=0}$$

erhalten, wobei natürlich $\sum_{i,k} g_{ik} c_i c_k = 1$ sein muß, wenn die Form (g_{ik}) auf die Koordinaten $\mathfrak{x}_\alpha$ bezogen ist. Bezeichnet man mit g_{ik}^0 den Wert von g_{ik} im Punkte (ξ), d. h. $\mathfrak{x}_\alpha = 0$, so wird insbesondere

$$\sum g_{ik}^0 c_i c_k = 1 \quad \text{oder} \quad s^2 = \sum g_{ik}^0 c_i c_k s^2 = \sum g_{ik}^0 \mathfrak{x}_i \mathfrak{x}_k,$$

d. h. die geodätische Entfernung ist eine quadratische Form in diesen Koordinaten mit *konstanten* Koeffizienten. Diese Beziehung braucht auch Hadamard bei seiner Konstruktion der Grundlösung und findet sie auf anderem Wege (a. a. O. S. 84–90). Herr Herglotz hat nun gezeigt (vgl. „Zur Riemannschen Metrik“, Berichte der sächsischen Akademie der Wiss. **73** (1921), S. 215; siehe auch Math. Annalen **93** (1925), S. 47), daß die Normalkoordinaten charakterisiert sind durch das Bestehen der n Identitäten

$$\sum g_{ik} c_k \equiv \sum g_{ik}^0 c_k \qquad (i = 1, \ldots, n),$$

woraus sich mühelos $\Delta_2\left(\frac{1}{s^{n-2}}\right) = -(n-2)\frac{1}{s^{n-1}}\frac{\partial \log \sqrt{g}}{\partial s}$ ergibt. Herr Herglotz setzt zwar die Koeffizienten analytisch voraus, doch dürfte das nicht notwendig sein. Unsere Behauptung kann man jedenfalls auch im allgemeinen Falle nachrechnen.

schränken, immer *positiv* voraussetzen, indem wir ihr eventuell eine Konstante hinzufügen. Um unsere Formeln möglichst den bekannten aus der Potentialtheorie anzupassen, können wir daher gelegentlich unter Auszeichnung eines Punktes (ξ) die Grundlösung in der Form

$$\Gamma(x;\xi) = \frac{1}{\varrho^{n-2}} \tag{12}$$

schreiben. Es ist dann ϱ beständig positiv, für $x_i \neq \xi_i$ regulär und es wird

$$\lim_{(\xi)\to(x)} \frac{s}{\varrho} = \lim_{s\to 0} \frac{s}{\varrho} = 1.$$

Die Hyperflächen $\varrho =$ konst. sind (für hinreichend kleine Werte der Konstanten) geschlossen, und speziell bestimmt die Gleichung $\varrho = 0$ nur den Punkt (ξ). Führt man die Größe ϱ und $n-1$ geeignete Parameter $\varphi_1, \ldots, \varphi_{n-1}$ auf den Hyperflächen $\varrho =$ konst. als neue Koordinaten ein, was innerhalb eines gewissen Bereiches immer möglich ist, so erhält das Linienelement die Form [7])

$$ds^2 = a_0\, d\varrho^2 + \sum_{\alpha,\beta=1}^{n-1} a_{\alpha\beta}\, d\varphi_\alpha\, d\varphi_\beta. \tag{13}$$

Das Koordinatensystem hat im Nullpunkte dieselbe Irregularität wie die gewöhnlichen Polarkoordinaten: die Diskriminante verschwindet wie ϱ^{2n-2}.

Das vorhin bestimmte *Oberflächenelement* berechnet sich für diese Flächen zu $\sqrt{a}\, d\varphi_1 \ldots d\varphi_{n-1}$, wenn a die Diskriminante der Form $\sum\limits_{\alpha,\beta=1}^{n-1} a_{\alpha\beta}\, d\varphi_\alpha\, d\varphi_\beta$ bedeutet (*nicht* des Linienelements; das Volumelement des Raumes ist demnach $\sqrt{a_0 a}\, d\varrho\, d\varphi_1 \ldots d\varphi_{n-1}$); die Differentiation in der Normalrichtung erweist sich gleich $\frac{1}{\sqrt{a_0}} \frac{\partial}{\partial \varrho}$. Es ist nun wichtig zu bemerken, daß der Ausdruck $\sqrt{\frac{a}{a_0}}$ die Größe ϱ nur als Faktor ϱ^{n-1} enthält, so daß $\frac{1}{\varrho^{n-1}} \sqrt{\frac{a}{a_0}} = \Phi(\varphi_1, \ldots, \varphi_{n-1})$ *von ϱ nicht mehr abhängt*. In unserem Koordinatensystem wird nämlich

$$\Delta_2 u = \frac{1}{\sqrt{a_0 a}} \left\{ \left(\sqrt{\frac{a}{a_0}}\, u_\varrho \right)_\varrho + \sum_{\alpha,\beta=1}^{n-1} \frac{\partial}{\partial \varphi_\beta} \left(a^{\alpha\beta} \sqrt{a_0 a}\, u_{\varphi_\alpha} \right) \right\}, \tag{14}$$

wobei jetzt auch die $a^{\alpha\beta}$ *bezüglich der* $(n-1)$-*reihigen Matrix* der $a_{\alpha\beta}$ gebildet sind. Da nun $u = \frac{1}{\varrho^{n-2}}$ eine Lösung sein muß, so folgt die Be-

[7]) Man kann immer die φ_i so wählen, daß die die Produkte enthaltenden „gemischten Glieder" verschwinden; vgl. z. B. Darboux, G., Théorie des Surfaces 2, 2. Aufl., (1915), S. 522.

hauptung unmittelbar. Daher ist insbesondere auch

$$\iint\limits_{\varrho=\text{konst.}} \frac{1}{\varrho^{n-1}}\sqrt{\frac{a}{a_0}}\,d\varphi_1\ldots d\varphi_{n-1} = E(\xi_1,\ldots,\xi_n) \tag{15}$$

eine von ϱ unabhängige beständig positive Größe.

§ 3.
Folgerungen. Der Mittelwertsatz.

Nun sind wir imstande diese Ergebnisse genau so wie in der Potentialtheorie anzuwenden und dadurch die bekannten Eigenschaften der Potentialfunktionen auf die Lösungen der allgemeinen selbstadjungierten elliptischen Differentialgleichung zu übertragen.

Setzen wir zunächst in die Greensche Formel $v = \Gamma(x;\xi) = \frac{1}{\varrho^{n-2}}$ ein, so müssen wir bei der Integration den Punkt $(\xi_1,\ldots,\xi_n)$, falls er sich innerhalb des Integrationsgebietes befindet, etwa durch eine Hyperfläche $\varrho = \varepsilon$ ausschließen. Der Beitrag dieser Hyperfläche zum $(n-1)$-fachen Integral auf der rechten Seite ist nach dem Gesagten (das negative Vorzeichen wegen der umgekehrten Orientierung)

$$-\iint\limits_{\varrho=\varepsilon}\left(\frac{u_\varrho}{\varrho^{n-2}} + (n-2)\frac{u}{\varrho^{n-1}}\right)\frac{1}{\sqrt{a_0}}\,do$$

$$= -\iint\limits_{\varrho=\varepsilon}\left(\frac{u_\varrho}{\varrho^{n-2}} + (n-2)\frac{u}{\varrho^{n-1}}\right)\sqrt{\frac{a}{a_0}}\,d\varphi_1\ldots d\varphi_{n-1}.$$

Da nun $\sqrt{\frac{a}{a_0}} = \varrho^{n-1}\,\Phi(\varphi_1,\ldots,\varphi_{n-1})$ ist, so verschwindet für $\varepsilon\to 0$ das Integral über den ersten Summanden, während das zweite in der Grenze $(n+2)\,E(\xi)\,u(\xi)$ ergibt. Es ist demnach für jeden Punkt (ξ) im Innern des Integrationsgebietes

$$(n+2)\,E(\xi)\,u(\xi) = -\iiint\limits_{G}\frac{1}{\varrho^{n-2}}\Delta_2 u\,dV + \iint\limits_{F}\left(\frac{u_n}{\varrho^{n-2}} - u\frac{\partial\frac{1}{\varrho^{n-2}}}{\partial n}\right)do. \tag{16}$$

Diese Formel stellt jede Lösung der Gleichung $\Delta_2 u = f(x_1,\ldots,x_n)$ dar als das „Potential" einer räumlichen „Massenverteilung" mit der Dichte f, und einer einfachen sowie einer „Doppelbelegung" auf dem Rande. Für Punkte (ξ) außerhalb des Integrationsgebietes wird natürlich

$$-\iiint\limits_{G}\frac{1}{\varrho^{n-2}}\Delta_2 u\,dV + \iint\limits_{F}\left(\frac{u_n}{\varrho^{n-2}} - u\frac{\partial\frac{1}{\varrho^{n-2}}}{\partial n}\right)do = 0; \tag{16 }$$

das ist wiederum die bekannte Sprungrelation, auf die sich aber alle übrigen zurückführen lassen.

Wir können diese Formel insbesondere auf Lösungen der Gleichung $\Delta_2 u = 0$ anwenden, so erhalten wir ein Analogon zum bekannten *Mittelwertsatz*, wenn wir für G das durch eine Hyperfläche $\varrho = \text{konst.} = A$ begrenzte Gebiet wählen und statt $\frac{1}{\varrho^{n-2}}$ die am Rande verschwindende Funktion $v = \frac{1}{\varrho^{n-2}} - \frac{1}{A^{n-2}}$ einsetzen. Wir erhalten

$$(n-2)\,E(\xi)\,u(\xi) = -\iint\limits_{\varrho=A} u\,\frac{\partial\frac{1}{\varrho^{n-2}}}{\partial n}\,do$$

oder ausgerechnet

$$u(\xi) = \frac{1}{E(\xi)}\iint\limits_{\varrho=A} u\,\frac{1}{\varrho^{n-1}\sqrt{a_0}}\,do = \frac{1}{E(\xi)}\iint\limits_{\varrho=A} u\,\Phi(\varphi_1,\ldots,\varphi_{n-1})\,d\varphi_1\ldots d\varphi_{n-1}. \tag{17}$$

Der Faktor von u unter dem Integrationszeichen ist beständig positiv und *hängt nicht von ϱ ab*; es ist weiter $E(\xi)\iint\limits_{\varrho=A}\Phi(\varphi_1,\ldots,\varphi_{n-1})\,d\varphi_1\ldots d\varphi_{n-1}$, so daß wir das genaue Analogon zum Mittelwertsatz vor uns haben, der den Funktionswert in einem Punkte (ξ) ausdrückt als Mittelwert der Funktionswerte auf einer beliebigen Hyperfläche Grundlösung = konst. mit diesem Punkt als Mittelpunkt. Nur ist hier die Funktion nicht bloß mit dem Oberflächenelement multipliziert, sondern noch mit dem Faktor $\frac{1}{\sqrt{a_0}}$, welcher eben bewirkt, daß das Integral unabhängig wird von der speziellen Fläche der Schar. Im euklidischen Falle ist $a_0 = 1$, und auch in nichteuklidischen Räumen mit konstantem Krümmungsmaß (sog. „projektiven Räumen") hängt a_0 bloß von ϱ ab; die Flächen $\varrho = \text{konst.}$ sind hier, wie man leicht nachrechnet, Kugeln, so daß die Analogie noch größer wird.

Selbstredend ergibt sich nun in bekannter Weise der Satz von der *Nichtexistenz eines Extremwertes* (im weiteren Sinne) usf. Im Falle von analytischen Koeffizienten folgt auch der *analytische Charakter* der Lösungen und die Möglichkeit der *analytischen Fortsetzung*. Wenn die Koeffizienten für *alle* Wertsysteme $x_1, \ldots, x_n$ den Differenzierbarkeitsannahmen genügen und beschränkt bleiben, so zwar, daß die Grundlösung im Unendlichen von der Ordnung $n-2$ verschwindet, so gilt auch das potentialtheoretische Analogon zum bekannten Satz von Liouville. Herr Sternberg[8]) hat übrigens gezeigt, daß man die Koeffizienten über einen gegebenen Bereich stets so fortsetzen kann, daß die Gleichung außerhalb einer gewissen Kugel mit der Laplaceschen Gleichung übereinstimmt. Für die so fortgesetzte Gleichung gilt der Satz immer.

---

[8]) A. a. O., S. 296–297.

Wir hätten auch von einer *nicht selbstadjungierten* Differentialgleichung

$$(8') \qquad \Delta_2 u + \sum_i b_i u_i + c u = 0$$

ausgehen können. Bezeichnet für den Moment $\frac{1}{\varrho^{n-2}}$ eine im Punkte (ξ) unendlich werdende Grundlösung der zu dieser Gleichung *adjungierten Gleichung*

$$(9') \qquad \Delta_2 v - \frac{1}{\sqrt{g}} \sum_i \frac{\partial}{\partial x_i}(b_i \sqrt{g}\, v) + c v = 0,$$

so erhalten wir wie vorhin gemäß (10)

$$(17') \qquad u(\xi) = \frac{1}{E'(\xi)} \iint\limits_{\varrho = A} u \frac{1}{\varrho^{n-1}\sqrt{a_0}}\, do,$$

da ja das Zusatzintegral $\iint\limits_{\varrho = A} N u \left(\frac{1}{\varrho^{n-2}} - \frac{1}{A^{n-2}}\right) do$ verschwindet. Natürlich ist jetzt der Faktor $\frac{1}{\varrho^{n-1}}\sqrt{\frac{a}{a_0}}$ nicht mehr von ϱ unabhängig, und es ist jetzt

$$(15') \qquad E'(\xi) = \lim_{\varepsilon \to 0} \iint\limits_{\varrho = \varepsilon} \frac{1}{\varrho^{n-1}\sqrt{a_0}}\, do.$$

Im übrigen ist zu bemerken, daß im Gegensatz zur Gleichung (17) diese Formel nicht unbeschränkt gültig ist, da jetzt die Grundlösung sehr wohl verschwinden kann und in diesem Falle die Transformation auf die neuen Koordinaten $\varrho, \varphi_1, \ldots, \varphi_{n-1}$ ungültig wird. Für hinreichend kleine Gebiete (jedenfalls solange die erste Randwertaufgabe lösbar ist) bleibt die Formel bestehen. Sie stellt den allgemeinsten Mittelwertsatz dar und umfaßt jenen von H. Weber[9]) für die spezielle Gleichung $\Delta u + K^2 u = 0$ ($K =$ konst.), bei welcher die Flächen $\varrho =$ konst. natürlich Kugeln sind und der Faktor $\frac{1}{\sqrt{a_0}}$ nur von ϱ abhängt.

Aus der Gleichung (17′) ergibt sich unmittelbar, daß eine Lösung u der Differentialgleichung (8′), die *in einem Bereiche ihr Vorzeichen nicht wechselt, daselbst auch nirgends verschwinden kann*, es sei denn, daß sie identisch Null ist; denn der Integrand wechselt ja in einer gewissen Umgebung eines jeden Punktes sein Vorzeichen nicht und kann daher auch nicht verschwinden. Daraus folgt nach einer Bemerkung von Herrn

---

[9]) Vgl. z. B. Encyklopädie der math. Wiss. 2 A, 7: Sommerfeld, Randwertaufgaben in der Theorie der partiellen Differentialgleichung, S. 541–542; oder auch H. Weber, Über die Integration der partiellen Differentialgleichung $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + K^2 u = 0$, Math. Annalen 1 (1869), S. 1–36.

Lichtenstein der Satz, daß eine (nicht konstante) Lösung dieser Differentialgleichung im Falle, daß $c = 0$, *weder ein Maximum noch ein Minimum (selbst im weiteren Sinne) haben kann.* Würde nämlich eine Lösung u in einem Punkte etwa einen Maximalwert M annehmen, so wäre die Funktion $M - u$, die ja auch eine Lösung der Differentialgleichung ist, in einer gewissen Umgebung des betreffenden Punktes nicht negativ, während der Punkt selbst eine Nullstelle dieser Funktion wäre. Das ist aber, wie eben gezeigt wurde, nicht möglich. Daraus folgen dann auch die Eindeutigkeitssätze für den Fall $c < 0$ und ähnliche Sätze.

§ 4.
Raum- und Oberflächenintegrale.

Wir kommen auf die selbstadjungierte Gleichung $\Delta_2 u = 0$ zurück. Man kann natürlich in Analogie zur Laplaceschen Gleichung „Potentiale" von räumlich oder flächenhaft verteilten „Massen" und von Doppelbelegungen betrachten. Das hat schon Herr Sternberg getan, und hat dabei die bekannten Sprungrelationen wiedergefunden. Er benutzt indessen ziemlich umständliche Abschätzungen, während man auch die äußerst einfachen Beweise von Herrn Erhard Schmidt übertragen kann. Die Analogie geht auch noch weiter, indem auch die Sätze über die analytische Fortsetzbarkeit gelten und im Sinne unserer allgemeinen Maßbestimmung sogar die *Gaußsche geometrische Deutung des „Potentials"* einer homogenen Doppelbelegung erhalten bleibt. Für den ersten Teil dieser Sätze übertragen sich die bekannten Sätze direkt, und es erübrigt sich daher näher darauf einzugehen; als Beispiel sei bloß die Poissonsche Gleichung hergeleitet.

Für unsere Potentiale benutzen wir zweckmäßig nicht eine beliebige Grundlösung, sondern normieren sie mit Herrn Sternberg so, daß das Integral (15) etwa immer gleich 1 wird, d. h. wir benutzen als Grundlösung $K(x;\xi) = \frac{\Gamma(x;\xi)}{E(\xi)}$. Von dieser Grundlösung hat Herr Sternberg[10]) die Symmetrie in x und ξ nachgewiesen, so daß sie auch als Funktion der ξ_i betrachtet der Differentialgleichung $\Delta_2 u$ genügt. Wir haben dann Integrale der Form

$$U(\xi) = \iiint_G \mu(x)\, K(x;\xi)\, dV, \tag{18}$$

$$V(\xi) = \iint_F \sigma(x)\, K(x;\xi)\, do, \tag{19}$$

$$W(\xi) = \iint_F \tau(x) \frac{\partial K(x;\xi)}{\partial n}\, do \tag{20}$$

[10]) A. a. O., S. 298.

zu betrachten, wobei μ, σ, τ mit hinreichenden Stetigkeitseigenschaften versehene Verteilungsfunktionen sind. Um etwa die Poissonsche Gleichung für (18) zu beweisen, betrachten wir mit Herrn E. Schmidt eine Hilfsfunktion v, die im Bereiche G der Differentialgleichung $\Delta_2 v = \mu$ genügt.

Es ist dann

$$U(\xi) = \iiint_G \Delta_2 v\, K(x;\xi)\, dV;$$

um die Greensche Formel anwenden zu können, müssen wir zunächst den Punkt (ξ) ausschließen und erhalten mit Rücksicht auf die Beziehungen (15) und $K(x;\xi) = \frac{\Gamma(x;\xi)}{E(\xi)}$ in bekannter Weise

$$u(\xi) = v(\xi) + \iint_F \left(v\frac{\partial K}{\partial n} - K\frac{\partial v}{\partial n}\right) do,$$

wenn F den Rand des Bereiches G bezeichnet. Da das Integral links eine reguläre Lösung der Differentialgleichung $\Delta_2 u = 0$ ist (wobei der Operator natürlich nach den Veränderlichen ξ_i genommen wird), so folgt ohne weiteres $\Delta_2 u = \Delta_2 v = \mu(\xi)$, und das ist die zu beweisende Poissonsche Gleichung.

Ebenso übertragen sich die Beweise für die anderen Sprungrelationen und verwandten Sätze.

Wir betrachten nun eine homogene Doppelschicht mit der Dichte 1:

$$W(\xi) = \iint_F \frac{\partial \Gamma(x;\xi)}{\partial n}\, do, \tag{21}$$

wobei F ein hinreichend stetiges, nicht notwendig geschlossenes Flächenstück ist, Γ eine *beliebige* Grundlösung bedeutet und der Punkt (ξ) außerhalb des Flächenstücks F gewählt wurde. Wir wählen als passendes Koordinatensystem etwa das in § 2 definierte und setzen voraus, daß das Flächenstück F von jeder Kurve $\varphi_\alpha =$ konst. $(\alpha = 1, \ldots, n-1)$ höchstens in einem Punkte geschnitten wird (sonst müßten wir das Flächenstück in einige Teilbereiche zerlegen und unsere Betrachtungen auf jeden einzelnen anwenden). Beschränken wir uns auf einen Bereich, in dem unsere Koordinaten regulär sind, so wird die Hyperfläche F durch die n Gleichungen

$$\varrho = \varrho(\lambda_1, \ldots, \lambda_{n-1}), \qquad \varphi_\alpha = \varphi_\alpha(\lambda_1, \ldots, \lambda_{n-1}) \quad (\alpha = 1, \ldots, n-1)$$

auf die Parameter $\lambda_1, \ldots, \lambda_{n-1}$ bezogen. Setzen wir noch

$$D_0 = \frac{\partial(\varphi_1, \ldots, \varphi_{n-1})}{\partial(\lambda_1, \ldots, \lambda_{n-1})} \quad \text{usf.},$$

so wird

$$\frac{\partial \Gamma}{\partial n} = \frac{\partial\left(\frac{1}{\varrho^{n-2}}\right)}{\partial n} = \frac{n-2}{\varrho^{n-1}}\frac{1}{a_0}\frac{D_0}{\sqrt{\frac{1}{a_0}D_0^2 + \sum\limits_{\alpha,\beta=1}^{n-1} a^{\alpha\beta} D_\alpha D_\beta}} = \frac{n-2}{\varrho^{n-1}}\cos(\varrho, n)$$

und

$$do = \sqrt{a_0 a}\sqrt{\frac{1}{a_0} D_0^2 + \sum_{\alpha,\beta=1}^{n-1} a^{\alpha\beta} D_\alpha D_\beta}\, d\lambda_1 \ldots d\lambda_{n-1}.$$

Dabei wurde die Normalenrichtung *umgekehrt*, so daß wir hier im Falle einer geschlossenen Fläche die *innere* Normalenrichtung gewählt haben. Daher wird

$$(22) \qquad W = (n-2)\iint \frac{1}{\varrho^{n-1}}\sqrt{\frac{a}{a_0}}\, D_0\, d\lambda_1 \ldots d\lambda_{n-1}$$
$$= (n-2)\iint \Phi(\varphi_1, \ldots, \varphi_{n-1})\, d\varphi_1 \ldots d\varphi_{n-1}.$$

Der Integrand hängt, wie im § 2 ausgeführt wurde, nicht von ϱ ab; das Integral wird über denjenigen Teil der Hyperflächen $\varrho =$ konst. erstreckt, auf welches das Flächenstück F durch die Kurven $\varphi_\alpha =$ konst. projiziert wird. Wir erhalten also das Potential der homogenen Doppelschicht als *ein Integral über den betreffenden Teil der Hyperfläche* $\varrho = 1$ (oder überhaupt $\varrho =$ konst.). Im Falle der Laplaceschen Gleichung ist das insbesondere die Einheitskugel, und da $a_0 = 1$ wird, ist der Integrand das Oberflächenelement, so daß das Potential durch den räumlichen Winkel, unter dem das Flächenstück F vom Punkte (ξ) aus gesehen wird, gegeben ist. Im allgemeinen Fall kommt überall das Oberflächenelement mit dem Faktor $\frac{1}{\sqrt{a_0}}$ multipliziert vor, welcher bewirkt, daß der Integrand nicht mehr von ϱ abhängt, und eben dadurch entspricht das Integral (22) vollkommen dem räumlichen Winkel. Speziell im „nichteuklidischen" Fall wo a_0 nur von ϱ abhängt, haben wir wiederum die Oberfläche der Einheitskugel multipliziert mit einer *reinen Zahl.*

Ist die Hyperfläche F geschlossen und befindet sich der Punkt (ξ) im Innern, so wird $W = (n-2)\, E(\xi)$, während im Äußeren $W = 0$ wird. Das ist die bekannte Sprungrelation. Hätten wir statt Γ die normierte Grundlösung K benutzt, so wäre selbstredend $E = 1$.

Man kann ebenso auch die allgemeinere Gaußsche Formel übertragen, die sich auf das Potential

$$U(\xi) = \iiint\limits_G \mu(x)\, K(x; \xi)\, dV$$

bezieht. Beachtet man nämlich, daß für alle Punkte innerhalb des Gebietes G $\Delta_2 U = \mu$ und außerhalb $\Delta_2 U = 0$ ist, so ergibt die Greensche Formel ähnlich wie bei der Laplaceschen Gleichung[11]), daß das Integral

$$\iint\limits_F \frac{\partial U}{\partial n}\, do,$$

[11]) Vgl. z. B. E. Goursat, Cours d'Analyse mathématique 3, S. 280 (3. Aufl. 1923).

(wobei nun nach ξ_i differenziert und integriert wird), erstreckt über eine beliebige, das Gebiet G nicht schneidende geschlossene Hyperfläche, *gleich ist der von dieser Fläche eingeschlossenen „Gesamtmasse"*, d. h. $\iiint\limits_G \mu\, dV$ bzw. Null. (Der übliche Faktor vom Betrage der Kugeloberfläche fehlt bei uns infolge der Normierung der Grundlösung.)

§ 5.

Sätze von Harnack.

Der erste Harnacksche Satz der Potentialtheorie besagt bekanntlich, daß eine Folge von in einem abgeschlossenen Bereiche regulären Potentialfunktionen gleichmäßig *konvergiert, falls die Randwerte es tun, und daß die Grenzfunktionen wiederum der Laplaceschen Gleichung genügt.* Dieser Satz folgt unmittelbar aus dem Satz vom Maximum und Minimum, und gilt daher ohne weiteres für die Lösungen der Gleichungen (8′), falls $c \leqq 0$. Herr Lichtenstein[12]) hat aber den Satz auch für die allgemeine Gleichung (8′) mit zwei unabhängigen Veränderlichen bewiesen, und sein Beweis kann wörtlich auf den Fall von n Veränderlichen übertragen werden, da er bloß den Satz für Potentialfunktionen und die Greensche Funktion benutzt.

Der zweite Harnacksche Satz lautet: *Wenn eine Folge* (regulärer) *nichtnegativer Potentialfunktionen in einem Punkte konvergiert, so konvergiert sie auch in jedem Punkte des Bereiches und zwar wiederum gegen eine Potentialfunktion.* Man beweist gewöhnlich diesen Satz, indem man aus der Poissonschen Formel durch Abschätzungen zum Hilfssatz gelangt: Sind P und Q zwei beliebige Punkte eines Kreises, so gibt es zwei Konstanten N und M, so daß in unmittelbar verständlicher Bezeichnung die Ungleichung

$$N u(Q) \leqq u(P) \leqq M u(Q) \tag{23}$$

gilt für *jede* in diesem Kreise reguläre und positive Potentialfunktion. Aus dieser Ungleichung ergibt sich unmittelbar eine eben solche für den ganzen Bereich, und daraus folgt dann der zu beweisende Satz leicht.

Wir sind nach dem Vorangegangenen imstande, ohne weiteres für die (positiven) Lösungen der *allgemeinsten* (homogenen) Differentialgleichung

$$\Delta_2 u + \sum b_i u_i + c u = 0 \tag{8'}$$

eine Ungleichung wie (23) herzuleiten, und dadurch den Harnackschen Satz zu übertragen.

---

[12]) L. Lichtenstein, Randwertaufgaben der Theorie der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus. I. Crelles Journ. **142** (1913), S. 1–40, insb. S. 14–15.

Betrachten wir einen willkürlich gewählten Punkt $P(\xi_1, \ldots, \xi_n)$ im Bereiche und die Schar der Hyperflächen $\Gamma(x; \xi) = \text{konst.}$ um diesen Punkt, wobei $\Gamma(x; \xi)$ die Grundlösung der *adjungierten* Differentialgleichung (9′) ist. Für eine hinreichend kleine Umgebung gilt der verallgemeinerte Mittelwertsatz (17′),

$$u(P) = \frac{1}{E'(P)} \iint\limits_{\varrho = A} u \frac{1}{\varrho^{n-1}\sqrt{a_0}}\, do,$$

wobei u eine beliebige Lösung der Differentialgleichung, und

$$do = \sqrt{a}\, d\varphi_1 \ldots d\varphi_{n-1}$$

ist. Andererseits erhalten wir, wenn wir mit $G(x; \mathfrak{x})$ die am Rande $\varrho = A$ verschwindende Greensche Funktion der adjungierten Differentialgleichung[18]) bezeichnen, aus den Greenschen Formeln in bekannter Weise den Funktionswert in einem beliebigen inneren Punkte $Q(\mathfrak{x}_1, \ldots, \mathfrak{x}_n)$

$$(24) \qquad \begin{aligned} u(Q) &= -\frac{1}{(n-2)E'(\mathfrak{x})} \iint\limits_{\varrho = A} u \frac{\partial G(x; \mathfrak{x})}{\partial n}\, do \\ &= -\frac{1}{(n-2)E'(\mathfrak{x})} \iint\limits_{\varrho = A} u \frac{\partial G(x; \mathfrak{x})}{\partial \varrho} \frac{1}{\sqrt{a_0}}\, do. \end{aligned}$$

Der Vergleich dieser beiden Formeln ergibt sofort, daß speziell für den Mittelpunkt ξ die Greensche Funktion

$$G(x; \xi) = -\frac{1}{A^{n-2}} + \frac{1}{\varrho^{n-2}}$$

ist. Nun ist für hinreichend kleine Werte A, auf die wir uns beschränkten, die Grundlösung und somit auch ϱ beständig positiv; dasselbe gilt von der Größe a_0 wegen dem elliptischen Charakter der Differentialgleichung $\left(\frac{1}{a_0}\right.$ ist ja der Koeffizient von u_ϱ in der Schreibweise von $\left.(14)\right)$. Es ist also der Faktor unter dem Integralzeichen (24) für $\mathfrak{x} = \xi$ positiv. Aus Stetigkeitsgründen gibt es daher eine ganze *abgeschlossene* Umgebung U des Punktes (ξ), so daß $-\frac{\partial G(x; \mathfrak{x})}{\partial n} = -\frac{1}{\sqrt{a_0}} \frac{\partial G}{\partial \varrho}$ für *alle* Punkte $(\mathfrak{x}_1, \ldots, \mathfrak{x}_n)$ dieser Umgebung und alle Punkte $(\xi_1, \ldots, \xi_n)$ der Hyperfläche $\varrho = A$ wesentlich positiv wird, und dasselbe gilt dann auch von $-\frac{1}{\varrho^{n-1}} \frac{\partial G}{\partial \varrho}$. Da nun die Hyperfläche $\varrho = A$ und ebenso die Umgebung U abgeschlossen sind, muß die Funktion ein *positives* Minimum μ und Maximum M haben.

---

[18]) Die Existenz dieser Greenschen Funktion ist klar, da wir uns auf hinreichend kleine Umgebungen beschränken.

Setzen wir auch von der Lösungsfunktion u voraus, daß sie nirgends negativ ist, so folgt aus dem Mittelwertsatz der Integralrechnung

$$(25) \qquad \frac{\mu}{E(\xi)}\iint\limits_{\varrho=A} u\,\frac{1}{\varrho^{n-1}\sqrt{a_0}}\,do \leqq u(\mathfrak{x}) \leqq \frac{M}{E(\xi)}\iint\limits_{\varrho=A} u\,\frac{1}{\varrho^{n-1}\sqrt{a_0}}\,do.$$

Das hier auftretende Integral ist das des Mittelwertsatzes (17′) und somit ergibt sich

$$(25') \qquad \mu u(\xi) \leqq u(\mathfrak{x}) \leqq M u(\xi);$$

dabei ist $(\mathfrak{x})$ ein *beliebiger* Punkt der Umgebung U, und die Konstanten μ und M hängen *nicht von der besonderen Wahl* der Funktion u ab. Dieselbe Ungleichung gilt für alle positiven Lösungsfunktionen u. Offenbar gilt auch für zwei *beliebige* Punkte $(\mathfrak{x})$ und $(\mathfrak{x}_1)$ des Bereiches U eine Ungleichung

$$\mu_1 u(\mathfrak{x}_1) \leqq u(\mathfrak{x}) \leqq M_1 u(\mathfrak{x})$$

mit $\mu_1 = \frac{\mu}{M}$, $M_1 = \frac{M}{\mu}$. Für *jeden* Punkt (ξ) kann man so eine Umgebung U finden. Nach dem bekannten Borelschen Überdeckungssatz genügen endlich viele dieser Umgebungen zur Überdeckung des ganzen Bereiches, und daraus folgert man, daß es für *jedes* Punktpaar (ξ) und $(\mathfrak{x})$ des Bereiches zwei Zahlen μ_0 und M_0 gibt, so daß die Ungleichung

$$(26) \qquad \mu_0 u(\xi) \leqq u(\mathfrak{x}) \leqq M_0 u(\xi)$$

gilt, für *jede* nicht negative Lösungsfunktion. Daraus folgt ohne weiteres, daß jede Folge solcher Lösungsfunktionen im ganzen Bereiche konvergiert (und zwar gleichmäßig), falls sie in einem Punkte (ξ) konvergiert. Aus der Gleichung (24) folgt weiter, daß die Grenzfunktion derselben Differentialgleichung genügt. Es ist also der *zweite Harnacksche Satz allgemein bewiesen für alle nicht negativen (d. h. aber nach* § 3 *positiven) Lösungsfunktionen der Differentialgleichung* (8′), selbst für Gebiete, für welche die Randwertaufgabe nicht lösbar ist. Im Falle zweier unabhängiger Veränderlicher wurde dieser Satz auf anderem Wege bereits von Herrn Lichtenstein bewiesen[14]).

Kiel, den 5. Februar 1929.

---

[14]) L. Lichtenstein, Beiträge zur Theorie der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus. Unendliche Folgen positiver Lösungen. Rendiconti del circolo mat. Palermo **33** (1912), S. 201–211.

(Eingegangen am 1. 3. 1929.)

Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben.

Von

Erich Rothe in Breslau.

Gegeben sei die in den Ableitungen der gesuchten Funktion z lineare Gleichung

$$\frac{\partial^2 z}{\partial y^2} = R(x, y)\frac{\partial z}{\partial x} + S(x, y, z). \tag{1}$$

Für $0 < y < 1$, $0 < x$ werde eine stetig nach x differenzierbare Lösung $z(x, y)$ von (1) gesucht, welche die Randbedingungen

$$z(0, y) = z_0(y); \quad z(x, 0) = 0; \quad z(x, 1) = 0 \tag{2}$$

erfüllt, wobei z_0 eine gegebene Funktion von y ist[1]).

---

[1]) Wegen der genaueren Voraussetzungen über die gegebenen Funktionen siehe S. 653, 656 und 666.

Die allgemeinere Aufgabe

$$\frac{\partial^2 z}{\partial y^2} = R(x, y)\frac{\partial z}{\partial x} + S(x, y, z) + T(x, y)\frac{\partial z}{\partial y} \qquad \left[\begin{array}{l}\chi_0(x) \neq \chi_1(x) \\ \chi_0(0) = 0 \\ \chi_1(0) = 1\end{array}\right]$$

$$z(0, y) = z_0(y); \quad z(x, \chi_0(x)) = f_0(x); \quad z(x, \chi_1(x)) = f_1(x)$$

läßt sich, wie leicht zu sehen, unter gewissen Differenzierbarkeitsvoraussetzungen über die gegebenen Funktionen $R, S, T, \chi_0, \chi_1, f_0, f_1$ auf das Problem (1), (2) zurückführen:

Zunächst erhält man durch die Transformation $\eta = \frac{y - \chi_0(x)}{\chi_1(x) - \chi_0(x)}$, $\xi = x$. $z(x, y) = u(\xi, \eta)$ die Gleichung

$$\frac{\partial^2 u}{\partial \eta^2} = R_1(\xi, \eta)\frac{\partial u}{\partial \xi} + S_1(\xi, \eta, u) + T_1(\xi, \eta)\frac{\partial u}{\partial \eta}$$

mit den Randbedingungen

$$u(0, \eta) = u_0(\eta); \quad u(\xi, 0) = f_0(\xi); \quad u(\xi, 1) = f_1(\xi).$$

Durch die Transformation

$$u = v e^{\frac{1}{2}\int_0^\eta T_1 d\eta}$$

(Fortsetzung der Fußnote [1]) auf nächster Seite.)

Wir wollen nun z in der folgenden Weise zu approximieren versuchen: wir teilen das Intervall von 0 bis X ($X > 0$) in n Teile der Länge $h = X:n$ und ersetzen (1), (2) durch

$$\frac{d^2 w_{\nu+1}}{dy^2} = R(x_{\nu+1}, y)\frac{w_{\nu+1} - w_\nu}{h} + S(x_{\nu+1}, y, w_\nu), \tag{1'}$$

$$w_0(y) = z_0(y); \quad w_{\nu+1}(0) = 0; \quad w_{\nu+1}(1) = 0. \tag{2'}$$

Da $w_0 = z_0$ gegeben ist, stellt (1'), (2') für $\nu = 0$ ein *ein*dimensionales lineares Randwertproblem für die Funktion $w_1(y)$ dar; allgemein stellt (1'), (2') eine eindimensionale lineare Randwertaufgabe für $w_{\nu+1}$ dar, wenn w_ν schon bekannt ist (und w_ν noch im Definitionsbereich des Koeffizienten S liegt).

Die Arbeit zerfällt in vier Teile: Im *ersten* werden zwei Hilfssätze über das Verhalten von Lösungen gewisser eindimensionaler Randwertprobleme zweiter Ordnung bei unendlich wachsendem, in der Differentialgleichung enthaltenem Parameter $\lambda = 1:h$ bewiesen. Im *zweiten* wird unter Voraussetzung der Existenz einer Lösung $z(x, y)$ des Problems (1), (2) gezeigt, daß die Differenz $z(x_\nu, y) - w_\nu(y)$ mit feiner werdender Teilung gegen Null strebt. Im *dritten* Teile wird gezeigt, wie man den Existenzbeweis für die Lösung von (1), (2) durch Grenzübergang aus den Lösungen $w_\nu(y)$ von (1'), (2') führen kann (wobei sich zugleich eine Abschätzung des Fehlers $z(x_\nu, y) - w_\nu(y)$ ergibt, S. 665)[2]. Bei diesem Beweis wird aber vorausgesetzt, daß (1) in den Punkten $x = 0$, $y = 0$ und $x = 0$, $y = 1$ erfüllt ist. Diese Voraussetzung zieht starke Einschränkungen für die gegebene Funktion $z_0(y)$ nach sich (S. 656, insbesondere Gleichung (24)).

---

erhält man eine Gleichung der Form

$$\frac{\partial^2 v}{\partial \eta^2} = R_2(\xi, \eta)\frac{\partial v}{\partial \xi} + S_2(\xi, \eta, v)$$

mit den Randbedingungen

$$v(0, \eta) = v_0(\eta); \quad v(\xi, 0) = g_0(\xi); \quad v(\xi, 1) = g_1(\xi).$$

Durch die weitere Transformation

$$\zeta(\xi, \eta) = v(\xi, \eta) - g_0(\xi)(1 - \eta) - g_1(\xi)\eta$$

erhält man schließlich ein Problem der Form (1), (2).

[2]) Wegen anderer Beweise für die Existenz von Lösungen des Problems vergleiche Gevrey, Journ. de Math. (6) 9 (1913) Equations aux dérivées partielles du type parabolique. — Die in der vorliegenden Arbeit befolgte Methode läßt sich auch auf gewisse Anfangswertprobleme partieller Differentialgleichungen anwenden. Die Anregung dazu, die von mir zunächst für Anfangswertprobleme durchgeführte Methode auf Randwertaufgaben zu übertragen, stammt von Herrn Prof. v. Mises. Vgl. auch dessen Vortrag: „Bemerkungen zur Hydrodynamik", Zeitschr. f. angew. Math. u. Mech. 7 (1927), insbesondere S. 439.

Im *vierten* Teil wird diese Einschränkung wieder aufgehoben, indem eine *beliebige* stetige (für $y=0$ und $y=1$ verschwindende) Funktion $z_0(y)$ durch Funktionen approximiert wird, die die Voraussetzungen des dritten Teils erfüllen, und die Konvergenz der zugehörigen Lösungen von (1) bewiesen wird. Hierbei wird benutzt, daß das Ergebnis des zweiten Teils ausreicht, um die Existenz einer Greenschen Funktion zu beweisen.

I.

Hilfssatz 1. Sei

$$w'' - \lambda\varrho(y)w = -\varrho(y)\varphi(y,\lambda), \tag{3}$$

$$w(0) = w(1) = 0 \tag{4}$$

ϱ und φ seien in $0 \leqq y \leqq 1$ stetige Funktionen von y und es sei $0 < m \leqq \varrho(y)$. φ sei überdies stetig in λ für $\lambda \geqq 0$. Bei festem $\lambda > 0$ ist dann

$$|w(y)| \leqq \frac{1}{\lambda}\,\mathrm{Max}\,|\varphi| \qquad (0 \leqq y \leqq 1). \tag{5}$$

Beweis. Nehmen wir zunächst $\varphi \geqq 0$ an. Dann ist auch $w \geqq 0$. Andernfalls hätte nämlich w in $(0,1)$ ein negatives Minimum. Da dieses wegen (4) nicht in $y=0$ oder $y=1$ angenommen werden kann, wäre an der betreffenden Stelle $w'' \geqq 0$, $\lambda\varrho w < 0$, $\varrho\varphi \geqq 0$, was nach (3) unmöglich ist. w nimmt daher ein positives Maximum an, und an der Stelle, wo dieses angenommen wird, ist nach (3)

$$\frac{\varphi}{\lambda} - w = -\frac{w''}{\lambda\varrho} \geqq 0.$$

Wegen $w \geqq 0$ folgt hieraus die Behauptung im Falle $\varphi \geqq 0$.

Zum Beweise des allgemeinen Falles bemerken wir, daß die Lösung von (3), (4) durch

$$w(y) = \int_0^1 \varrho(\eta)\,\varphi(\eta,\lambda)\,g(y,\eta,\lambda)\,d\eta \tag{6}$$

gegeben wird, wenn g die zu der Randbedingung (4) gehörige Greensche Funktion von $w'' - \lambda\varrho w$ ist. Nun ist $g(y,\eta,\lambda) \geqq 0$, wenn y und η in $(0,1)$ liegen. Daher folgt aus (6)

$$|w(y)| \leqq \int_0^1 \varrho(\eta)\,|\varphi(\eta,\lambda)|\,g(y,\eta,\lambda)\,d\eta. \tag{7}$$

Das rechts stehende Integral ist aber die Lösung von (3), (4), wenn in (3) φ durch $|\varphi|$ ersetzt wird und daher nach dem schon Bewiesenen gewiß $\leqq \frac{\mathrm{Max}\,|\varphi|}{\lambda}$, so daß aus (7) die zu beweisende Ungleichung (5) folgt.

Hilfssatz 2. Sei $w(y)$ die durch

$$w(0) = \alpha, \quad w(1) = \beta$$

bestimmte Lösung von (3). Dann ist bei festem $\lambda > 0$

$$(8) \qquad |w| \leqq \frac{1}{\lambda} \operatorname{Max} |\varphi| + \operatorname{Max}(|\alpha|, |\beta|) \qquad (0 \leqq y \leqq 1).$$

Beweis. Wir setzen $w = w_1 + w_2$, wo w_1 und w_2 durch

$$w_1'' - \lambda \varrho(y) w_1 = -\varrho(y)\varphi(y, \lambda) \qquad w_1(0) = w_1(1) = 0$$
$$w_2'' - \lambda \varrho(y) w_2 = 0 \qquad w_2(0) = \alpha;\ w_2(1) = \beta$$

bestimmt sind. Da w_2 in einem *inneren* Punkte von $(0, 1)$ weder ein positives Maximum noch ein negatives Minimum annehmen kann, ist $|w_2| \leqq \operatorname{Max}(|\alpha|, |\beta|)$; daher folgt die Behauptung (8) aus Hilfssatz 1.

II.

Satz 1. *Im Bereiche*

$$(9) \qquad 0 \leqq x \leqq X, \quad 0 \leqq y \leqq 1, \quad |z| \leqq Z$$

seien die Koeffizienten R, S von (1) *nebst ihren ersten Ableitungen nach z stetige Funktionen ihrer Argumente; ferner sei R einmal stetig nach x, zweimal stetig nach y differenzierbar und es sei im Bereiche* (9)

$$(10) \qquad R(x, y) \geqq m > 0.$$

Wir behaupten: Wenn eine Lösung $z(x, y)$ von (1), (2) *existiert, die in*

$$(11) \qquad 0 \leqq x \leqq a < X, \quad 0 \leqq y \leqq 1$$

nebst ihrer ersten Ableitung nach x stetig ist (auch am Rande des Bereiches (11)) *und dort der Bedingung*

$$(12) \qquad |z| \leqq c < Z$$

genügt, so gibt es zu jedem positiven $\varepsilon < Z - c$ eine positive Zahl $n_0 = n_0(\varepsilon)$, so daß für $n > n_0$

$$(13) \qquad |z(x_\nu, y) - w_\nu(y)| < \varepsilon \qquad \left(x_\nu = \frac{\nu X}{n};\ \nu = 0, 1, \ldots \left[\frac{na}{X}\right]\right),$$

wo $w_\nu(y)$ die durch (1′), (2′) *bestimmten Funktionen sind.*

Beweis. Setzen wir $z(x_\nu, y) = z_\nu(y)$ und

$$(14) \qquad \left(\frac{\partial z}{\partial x}\right)_{x = x_{\nu+1}} - \frac{z_{\nu+1} - z_\nu}{h} = \eta_{\nu+1}(y) \qquad \left(h = \frac{X}{n};\ \nu + 1 = 1, 2, \ldots \left[\frac{a}{h}\right]\right),$$

so liefert (1) für $x = x_{\nu+1}$

$$(15) \qquad \frac{d^2 z_{\nu+1}}{dy^2} = R(x_{\nu+1}, y)\frac{z_{\nu+1} - z_\nu}{h} + S(x_{\nu+1}, y, z_{\nu+1}) + R(x_{\nu+1}, y)\,\eta_{\nu+1}(y).$$

Rechnen wir zunächst *formal*[3]) und ziehen (1′) von (15) ab, so erhalten wir, wenn noch

(16) $$z_{\nu+1}(y) - w_{\nu+1}(y) = \gamma_{\nu+1}(y)$$

gesetzt wird,

$$\frac{d^2\gamma_{\nu+1}}{dy^2} = R(x_{\nu+1}, y)\frac{\gamma_{\nu+1}-\gamma_\nu}{h} + S(x_{\nu+1}, y, z_{\nu+1}) - S(x_{\nu+1}, y, w_\nu) + R(x_{\nu+1}, y)\eta_{\nu+1}$$
$$= R(x_{\nu+1}, y)\frac{\gamma_{\nu+1}-\gamma_\nu}{h} + \frac{\overline{\partial S}}{\partial z}(z_{\nu+1} - z_\nu + \gamma_\nu) + R(x_{\nu+1}, y)\eta_{\nu+1}$$

(dabei bedeutet das Überstreichen, daß für das Argument z gewisse Mittelwerte zwischen $z_{\nu+1}$ und w_ν zu nehmen sind). $\gamma_{\nu+1}$ genügt also der Differentialgleichung

(17) $$\gamma''_{\nu+1} - \lambda\varrho\gamma_{\nu+1} = \varrho(y)\,\psi(y, \lambda),$$

worin

$$\lambda = \frac{1}{h}, \quad \varrho(y) = R(x_{\nu+1}, y)$$

und

(18) $$\psi(y, \lambda) = -\lambda\gamma_\nu + \frac{\overline{\partial S}}{\partial z}\frac{z_{\nu+1} - z_\nu + \gamma_\nu}{R(x_{\nu+1}, y)} + \eta_{\nu+1}$$
$$= -\lambda\gamma_\nu\left\{1 - \frac{1}{\lambda R(x_{\nu+1}, y)}\frac{\overline{\partial S}}{\partial z}\right\} + \frac{\overline{\partial S}}{\partial z}\frac{z_{\nu+1} - z_\nu}{R(x_{\nu+1}, y)} + \eta_{\nu+1}$$

gesetzt ist. Außerdem ist wegen (2), (2′) und (16)

(19) $$\gamma_{\nu+1}(0) = \gamma_{\nu+1}(1) = 0 \quad \text{und} \quad \gamma_0(y) = 0.$$

Unter der Annahme, daß w_ν stetig und absolut $< Z$ (deren Richtigkeit noch zu beweisen ist), genügt daher das Problem (17), (19) erstens allen Voraussetzungen des Hilfssatzes 1, und zweitens ist auf Grund der gemachten Voraussetzungen nach (18)

$$|\psi(y, \lambda)| \leq \lambda|\gamma_\nu|\left(1 + \frac{A}{\lambda}\right) + \frac{A}{2}\left(\frac{1}{\lambda} + \tau_1\right),$$

wo τ_1 gleichmäßig in (49) mit h gegen Null geht, und A eine passend gewählte, von y, ν und λ unabhängige positive Konstante ist. Aus Hilfssatz 1 folgt daher

(20) $$|\gamma_{\nu+1}(y)| \leq \frac{1}{\lambda}\operatorname{Max}|\psi(y, \lambda)| \leq \operatorname{Max}|\gamma_\nu|(1 + Ah) + \frac{1}{2}A(h + \tau_1)h.$$

Zu *beweisen* ist nun erstens, daß die Annahme $|w_\nu| < Z$ und w_ν stetig berechtigt ist, und zweitens, daß γ_ν für $\nu = 0, 1, \ldots, \left[\frac{a}{h}\right]$ mit h gleichmäßig

---

[3]) Gleichung (1′) hat nur einen Sinn, wenn schon gezeigt ist, daß w_ν noch im Definitionsbereich von S liegt, d. h. daß $|w_\nu| < Z$ ist.

in y und ν gegen Null geht. Hierzu wollen wir induktiv die folgenden beiden Behauptungen beweisen:

τ sei die nicht kleinere der Zahlen h und τ_1. Da dann τ gleichmäßig in x und y mit h gegen Null geht, gibt es eine Zahl h_0, so daß für $0 < h \leqq h_0$

$$(21) \qquad \tau \cdot A X e^{AX} < Z - c$$

ist (vgl. (12)). Für $0 < h < h_0$ behaupten wir nun:

a) $w_{\nu+1}$ ist stetig und $|w_{\nu+1}| < Z$ $\left(\nu = -1, 0, 1, \ldots, \left[\frac{a}{h}\right] - 1\right)$,

b) $|\gamma_{\nu+1}| \leqq (\nu + 1) h \tau (1 + Ah)^\nu A$.

Für $\nu = -1$ sind diese Behauptungen richtig: die Behauptung a) wegen $w_0 = z_0 = z(0, y)$ nach (12), die Behauptung b) wegen. $\gamma_0 = 0$. Nehmen wir nun an, daß a) und b) richtig seien, wenn man ν durch $\nu - 1$ ersetzt. Dann folgt wegen a), daß die Argumente der in (1′) auftretenden Funktionen innerhalb des Bereiches (9) liegen. Die frühere formale Schlußweise besteht dann zu Recht und es gilt die Ungleichung (20). Aus dieser folgt auf Grund der Induktionsannahme b) (in der $\nu - 1$ an Stelle von ν zu setzen ist):

$$|\gamma_{\nu+1}| \leqq \nu h \tau (1 + Ah)^{\nu-1} A (1 + Ah) + \frac{1}{2} A (h + \tau_1) h$$
$$\leqq \nu h \tau (1 + Ah)^\nu A + A h \tau < \nu \cdot h \tau (1 + Ah)^\nu A + h \tau (1 + Ah)^\nu A,$$

womit die Behauptung b) erwiesen ist. a) ist aber ebenfalls richtig; denn es ist erstens $w_{\nu+1}$ als Lösung des Problems (1′), (2′) gewiß stetig und zweitens nach (12), b) und (21)

$$|w_{\nu+1}| \leqq |z_{\nu+1}| + |\gamma_{\nu+1}| \leqq c + (\nu + 1) h \tau (1 + Ah)^\nu A$$
$$< c + X\tau \left(1 + \frac{AX}{n}\right)^n A < c + \tau A X e^{AX} < Z.$$

Hiermit sind a) und b) allgemein bewiesen. Nunmehr folgt leicht, daß γ_ν gleichmäßig in y und ν gegen Null geht, denn es ist nach b):

$$(22) \qquad |\gamma_{\nu+1}| \leqq a \tau (1 + Ah)^\nu A \leqq a \tau A \left(1 + \frac{AX}{n}\right)^{\left[\frac{na}{X}\right]}$$
$$\leqq a \tau A \left(1 + \frac{XA}{n}\right)^{\frac{na}{X}} < a \tau A e^{aA}.$$

Hiermit ist Satz 1 bewiesen, da τ gleichmäßig in x und y mit h gegen Null geht.

Folgerung. Aus dem eben bewiesenen Satze ergibt sich unmittelbar, daß es höchstens eine Lösung des Problems (1), (2) gibt, welche die in dem Satze angeführten Voraussetzungen erfüllt.

III.

Satz 2. *Vorgelegt sei das Problem* (1), (2). *Die Koeffizienten von* (1) *mögen den in Satz* 1 *gemachten Voraussetzungen genügen. Die gegebene Funktion* $z_0(y)$ *sei viermal stetig differenzierbar und es sei*

$$(23)\qquad z_0(0)=0;\quad z_0(1)=0;\quad |z_0(y)|<Z \qquad \text{für } 0\leqq y\leqq 1.$$

Außerdem werde vorausgesetzt, daß die Differentialgleichung (1) *in den Punkten* (0, 0) *und* (0, 1) *des Randes erfüllt sei, d. h. es sei wegen* (1), (2):

$$(24)\qquad \left(\frac{d^2 z_0}{dy^2}\right)_{y=0}=S(0,0,z_0(0));\qquad \left(\frac{d^2 z_0}{dy^2}\right)_{y=1}=S(0,1,z_0(1)).\,^{4)}$$

Ferner sollen $\frac{\partial S}{\partial x}, \frac{\partial^2 S}{\partial x^2}, \frac{\partial S}{\partial y}, \frac{\partial^2 S}{\partial y^2}, \frac{\partial S}{\partial z}, \frac{\partial^2 S}{\partial z^2}, \frac{\partial^2 R}{\partial x^2}$ *und* $\frac{\partial^2 S}{\partial x\,\partial z}$ *existieren und im Bereiche* (9) *stetig sein. Dann gibt es bei passender Wahl der Konstanten* a *eine* (*und, wie aus der am Schluß von* II (S. 655) *gemachten Bemerkung folgt, nur eine*) *Lösung von* (1), (2), *die im Bereiche* (11) *einmal stetig nach* x *differenzierbar ist.*

Beweis. Wir gehen aus von dem Problem (1′), (2′) und werden zunächst einige Eigenschaften der Lösungen $w_\nu(y)$ dieses Problems feststellen.

1. Es gibt eine von ν, y und h unabhängige Konstante L, so daß

$$(25)\qquad |w_{\nu+1}(y)-w_\nu(y)|<Lh \qquad (0\leqq y\leqq 1)$$

für alle $\nu=0,1,2,\ldots$ ist, die der Ungleichung

$$(26)\qquad x_\nu=\nu h\leqq a-h$$

genügen, wobei für die positive Konstante a die Ungleichungen

$$(27)\qquad a<X,$$

$$(28)\qquad a<(Z-|z_0|)\frac{e^{-KX}}{(1+X)K}$$

---

4) Verlangt man, daß die Lösung z sowie $\frac{\partial z}{\partial x}$ stetig auch am *Rande* des Bereiches (11) sind, so ist die Bedingung (24) [und natürlich auch (23)], wie man sich leicht überlegt, *notwendig*. [Implizite ist (23), (24) also schon bei Satz 1 vorausgesetzt.] Bei der hier befolgten Methode, $\frac{\partial z}{\partial x}$ durch Differenzenquotienten zu approximieren, wobei von $x=0$ ausgegangen wird, erscheint die Voraussetzung der Stetigkeit von $\frac{\partial z}{\partial x}$ auch für $x=0$ jedenfalls als naheliegend. Die Befreiung von der Bedingung (24) erfolgt in IV.

gelten. Die von ν, y und h unabhängige Zahl K soll dabei den Ungleichungen

$$\frac{1}{m}\left|S - z_0''\right| < K, \tag{29}$$

$$\frac{1}{m}\left|\frac{\partial R}{\partial x} + \frac{\partial S}{\partial z}\right| < K, \tag{30}$$

$$\frac{1}{m}\left|\frac{\partial S}{\partial x}\right| < K \tag{31}$$

genügen. Um nun (25) zu beweisen, werden wir zeigen: Für die der Ungleichung (26) genügenden ν ist

$$|w_{\nu+1} - w_\nu| < Kh(1+Kh)^\nu(1+h\nu). \tag{32}$$

Aus (32) folgt in der Tat unsere Behauptung (25), denn es ist wegen $h = X:n$

$$(1+Kh)^\nu = \left(1+\frac{KX}{n}\right)^\nu \leqq \left(1+\frac{KX}{n}\right)^n < e^{KX}, \qquad h\nu = x_\nu \leqq X, \tag{33}$$

so daß (25) mit

$$L = Ke^{KX}(1+X) \tag{34}$$

erfüllt ist.

(32) beweisen wir induktiv. Aus (1′) für $\nu = 0$ folgt durch Subtraktion von $w_0'' = z_0''$

$$(\varDelta w_0)'' - R(x_1, y)\frac{\varDelta w_0}{h} = S(x_1, y, z_0) - z_0'' = R(x_1, y)\frac{S(x_1, y, z_0) - z_0''}{R(x_1, y)} \text{ }^{5})$$

und aus (2′), (23) folgt

$$(\varDelta w_0)_{y=0} = (\varDelta w_0)_{y=1} = 0.$$

Aus Hilfssatz 1 in Verbindung mit (29) folgt daher $|\varDelta w_0| < hK$. Hiermit ist (32) für $\nu = 0$ bewiesen.

Um die Ungleichung, die aus (32) entsteht, wenn ν durch $\nu+1$ ersetzt wird (unter Annahme der Gültigkeit von (32) für $\nu = 0, 1, \ldots, \nu$), zu beweisen, falls noch $(\nu+1)h \leqq a - h$, bemerken wir zunächst, daß unter Beachtung von (33) aus der Induktionsannahme

$$\begin{aligned} |w_{p+1} - w_0| &\leqq |\varDelta w_0| + |\varDelta w_1| + \ldots + |\varDelta w_p| < (p+1)hKe^{KX}(1+X) \\ &< aKe^{KX}(1+X) \qquad (p = 0, 1, \ldots, \nu), \end{aligned}$$

also wegen $w_0 = z_0$

$$|w_{p+1} - z_0| < aKe^{KX}(1+X) \qquad (p = 0, 1, \ldots, \nu)$$

oder nach (28)

$$|w_{p+1}| < Z \qquad (p = 0, 1, \ldots, \nu)$$

---

[5]) Das Zeichen $\varDelta$ ist hier wie im folgenden in bezug auf ν bzw. x, nicht in bezug auf y gemeint. Also:

$$\varDelta w_\nu = w_{\nu+1}(y) - w_\nu(y); \quad \varDelta^2 w_\nu = w_{\nu+2}(y) - 2w_{\nu+1}(y) + w_\nu(y).$$

folgt. Daher liegen $x_{\nu+1}, y, w_\nu$ und $x_{\nu+2}, y, w_{\nu+1}$ noch im Bereiche $0 \leqq x \leqq a$, $0 \leqq y \leqq 1$, $|z| < Z$. Ziehen wir nun (1'), (2') von denjenigen Gleichungen ab, die entstehen, wenn ν durch $\nu+1$ ersetzt wird:

$$(\Delta w_{\nu+1})'' - R(x_{\nu+2}, y)\frac{\Delta w_{\nu+1}}{h}$$

$$= -R(x_{\nu+1}, y)\frac{\Delta w_\nu}{h} + S(x_{\nu+2}, y, w_{\nu+1}) - S(x_{\nu+1}, y, w_\nu)$$

$$= -\frac{R(x_{\nu+2}, y)}{h}\left\{\Delta w_\nu\left[1 + \frac{R(x_{\nu+1}, y) - R(x_{\nu+2}, y)}{R(x_{\nu+2}, y)}\right]\right.$$

$$\left. - h\frac{S(x_{\nu+2}, y, w_{\nu+1}) - S(x_{\nu+1}, y, w_\nu)}{R(x_{\nu+2}, y)}\right\},$$

$$(\Delta w_{\nu+1})'' - R(x_{\nu+2}, y)\frac{\Delta w_{\nu+1}}{h}$$

$$= -\frac{R(x_{\nu+2}, y)}{h}\left\{\Delta w_\nu\left[1 - \frac{h}{R(x_{\nu+2}, y)}\left(\frac{\partial R}{\partial x} + \frac{\partial S}{\partial z}\right)\right] - \frac{h^2}{R(x_{\nu+2}, y)}\frac{\partial S}{\partial x}\right\}^{*)},$$

$$(\Delta w_{\nu+1})_{y=0} = (\Delta w_{\nu+1})_{y=1} = 0.$$

Auf Grund von Hilfssatz 1 ist nun $|\Delta w_{\nu+1}|$ kleiner als das Maximum des Betrages des in der geschweiften Klammer stehenden Ausdrucks. Daher wird nach (30), (31)

$$|\Delta w_{\nu+1}| \leqq \mathrm{Max}\{|\Delta w_\nu|\}(1 + Kh) + Kh^2,$$

also nach Induktionsannahme (32)

$$\begin{aligned} |\Delta w_{\nu+1}| &< Kh(1+Kh)^{\nu+1}(1+h\nu) + Kh^2 \\ &< Kh(1+Kh)^{\nu+1}(1+h\nu) + Kh^2(1+Kh)^{\nu+1} \\ &= Kh(1+Kh)^{\nu+1}[1 + h(\nu+1)], \end{aligned}$$

womit (32) allgemein und damit (25) bewiesen ist.

2. Es gibt eine von y und h unabhängige Konstante L_1, so daß

$$(\Delta w_0)'' < L_1 h \tag{35}$$

ist. Zum Beweise subtrahieren wir w_0'' von Gleichung (1') für $\nu = 0$:

$$(\Delta w_0)'' - R(x_1, y)\frac{\Delta w_0}{h} = S(x_1, y, w_0) - w_0''. \tag{36}$$

Addieren wir $h\left[\frac{S(x_1, y, w_0) - w_0''}{R(x_1, y)}\right]''$, so erhalten wir

$$\left[\Delta w_0 + h\frac{S(x_1, y, w_0) - w_0''}{R(x_1, y)}\right]'' - \frac{R(x_1, y)}{h}\left[\Delta w_0 + h\frac{S(x_1, y, w_0) - w_0''}{R(x_1, y)}\right] \tag{37}$$

$$= h\left(\frac{S(x_1, y, w_0) - w_0''}{R(x_1, y)}\right)''.$$

---

*) Die nicht näher bezeichneten Argumente von $\frac{\partial R}{\partial x}, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial z}$ sind für x bzw. z gewisse Mittelwerte zwischen $x_{\nu+1}$ und $x_{\nu+2}$ bzw. w_ν und $w_{\nu+1}$.

Für $y = 0$ wird nun nach (2′), (23), (24) und (10)

$$\left|\Delta w_0 + h\frac{S(x_1,0,w_0)-w_0''}{R(x_1,0)}\right| = h\left|\frac{S(x_1,0,0)-S(0,0,0)}{R(x_1,0)}\right| < \frac{h^2}{m}\operatorname{Max}\left|\frac{\partial S}{\partial x}\right|$$

und die gleiche Abschätzung gilt für $y = 1$. Aus (37) folgt daher nach Hilfssatz 2

$$\left|\Delta w_0 + h\frac{S(x_1,y,w_0)-w_0''}{R(x_1,y)}\right| \leqq \frac{h^2}{m}\left\{\operatorname{Max}\left|\left(\frac{S(x,y,w_0)-w_0''}{R}\right)''\right| + \operatorname{Max}\left(\frac{\partial S}{\partial x}\right)\right\}.$$

Hieraus folgt nach (36) die Behauptung (35) mit

$$L_1 = \frac{\operatorname{Max} R}{m}\left[\operatorname{Max}\left|\left(\frac{S(x,y,w_0)-w_0''}{R}\right)''\right| + \operatorname{Max}\left|\frac{\partial S}{\partial x}\right|\right].$$

3. Es gibt eine von ν, h und y unabhängige Konstante L_2, so daß

(38) $$|\Delta^2 w_\nu| < L_2 h^2 \qquad (\text{für } (\nu+2)h < a).$$

Zum Beweise wählen wir eine Konstante K_1, so daß

(39) $$\frac{1}{m}\left[L\operatorname{Max}\left|\frac{\partial R}{\partial x}\right| + \operatorname{Max}\left|\frac{\partial S}{\partial x}\right| + L\operatorname{Max}\left|\frac{\partial S}{\partial z}\right| + L_1\right] < K_1,$$

(40) $$\frac{1}{m}\left[2\operatorname{Max}\left|\frac{\partial R}{\partial x}\right| + \operatorname{Max}\left|\frac{\partial S}{\partial z}\right|\right] < K_1,$$

(41) $$\frac{1}{m}\left[L\operatorname{Max}\left|\frac{\partial^2 R}{\partial x^2}\right| + \operatorname{Max}\left|\frac{\partial^2 S}{\partial x^2}\right| + 2L\operatorname{Max}\left|\frac{\partial^2 S}{\partial x\,\partial z}\right| + L^2\operatorname{Max}\left|\frac{\partial^2 S}{\partial z^2}\right| < K_1.$$

Dann behaupten wir

(42) $$|\Delta^2 w_\nu| < K_1 h^2 (1+K_1 h)^\nu (1+h\nu) \qquad ((\nu+2)h \leqq a).$$

Hieraus folgt die Behauptung (38) mit $L_2 = K_1 e^{K_1 X}(1+X)$ (vgl. den entsprechenden Schluß von (32) auf (25)).

Um nun (42) für $\nu = 0$ zu beweisen, ziehen wir von Gleichung (1′) für $\nu = 0$ auf beiden Seiten w_0'' und ebenso w_1'' von Gleichung (1′) für $\nu = 1$ ab. Die so entstandenen Gleichungen subtrahieren wir voneinander und erhalten

$$(w_2 - w_1)'' - (w_1 - w_0)'' = R(x_2,y)\frac{w_2 - w_1}{h} - R(x_1,y)\frac{w_1 - w_0}{h}$$
$$+ S(x_2,y,w_1) - S(x_1,y,w_0) - (w_1'' - w_0'')$$

oder

(43) $$(\Delta^2 w_0)'' - R(x_2,y)\frac{\Delta^2 w_0}{h}$$
$$= R(x_2,y)\left[\frac{w_1-w_0}{h}\,\frac{R(x_2,y)-R(x_1,y)}{R(x_2,y)} + \frac{1}{R(x_2,y)}\left(h\frac{\partial S}{\partial x} + \Delta w_0\frac{\partial S}{\partial z} - (\Delta w_0)''\right)\right]$$

(Argument von $\frac{\partial S}{\partial x}$ und $\frac{\partial S}{\partial z}$ in bezug auf x bzw. z Zwischenwerte zwischen x_1, x_2 bzw. w_0, w_1). Da außerdem nach (2′) und (23)

$$(\Delta^2 w_0)_{y=0} = (\Delta^2 w_0)_{y=1} = 0,$$

so folgt nach Hilfssatz 1, daß $|\Delta^2 w_0|:h$ kleiner als das Maximum des absolut genommenen in der eckigen Klammer der Gleichung (43) stehenden Ausdrucks ist. Dieser ist aber nach (25), (10) und (35) dem Betrage nach kleiner als

$$\frac{h}{m}\left[L \operatorname{Max}\left|\frac{\partial R}{\partial x}\right| + \operatorname{Max}\left|\frac{\partial S}{\partial x}\right| + L \operatorname{Max}\left|\frac{\partial S}{\partial z}\right| + L_1\right].$$

Hieraus im Verein mit (39) folgt (42) für $\nu = 0$. Wir wollen jetzt (42) für $\nu + 1$ unter Annahme der Gültigkeit für ν beweisen, falls noch $(\nu + 3) h \leqq a$. Geeignete Kombination dreier aufeinanderfolgender der Gleichungen (1′) liefert für $(\nu + 3) h \leqq a$

$$\frac{d^2 \Delta^2 w_{\nu+1}}{dy^2} = \Delta^2 \left\{R(x_{\nu+1}, y)\frac{w_{\nu+1} - w_\nu}{h}\right\} + \Delta^2 S(x_{\nu+1}, y, w_\nu)$$
$$= R(x_{\nu+3}, y)\frac{\Delta^2 w_{\nu+1} - \Delta^2 w_\nu}{h} + [R(x_{\nu+3}, y) - R(x_{\nu+1}, y)]\frac{\Delta^2 w_\nu}{h}$$
$$+ \frac{\Delta w_{\nu+1}}{h}\Delta^2 R(x_{\nu+1}, y) + \Delta^2 S(x_{\nu+1}, y, w_\nu),$$

oder

$$\text{(44)} \qquad \frac{d^2 \Delta^2 w_{\nu+1}}{dy^2} - R(x_{\nu+3}, y)\frac{\Delta^2 w_{\nu+1}}{h} = -\frac{R(x_{\nu+3}, y)}{h}$$
$$\left\{\Delta^2 w_\nu\left[1 - \frac{R(x_{\nu+3}, y) - R(x_{\nu+1}, y)}{R(x_{\nu+3}, y)}\right] - \frac{\Delta w_{\nu+1}\cdot\Delta^2 R(x_{\nu+1}, y) + h\Delta^2 S(x_{\nu+1}, y, w_\nu)}{R(x_{\nu+3}, y)}\right\}.$$

Außerdem ist nach (2′)

$$\text{(45)} \qquad (\Delta^2 w_{\nu+1})_{y=0} = (\Delta^2 w_{\nu+1})_{y=1} = 0.$$

Wir können daher wieder Hilfssatz 1 anwenden. Beachten wir, daß

$$\Delta^2 S(x_{\nu+1}, y, w_\nu) = S(x_{\nu+3}, y, w_{\nu+2} - 2 S(x_{\nu+2}, y, w_{\nu+1}) + S(x_{\nu+1}, y, w_\nu)$$
$$= S(x_{\nu+1} + 2h, y, w_\nu + \Delta w_\nu + \Delta w_{\nu+1}) - S(x_{\nu+1} + h, y, w_\nu + \Delta w_\nu)$$
$$- S(x_{\nu+1} + h, y, w_\nu + \Delta w_{\nu+1}) + S(x_{\nu+1}, y, w_\nu)$$
$$+ S(x_{\nu+1} + h, y, w_\nu + \Delta w_{\nu+1}) - S(x_{\nu+1} + h, y, w_\nu + \Delta w_\nu),$$

also

$$\Delta^2 S(x_{\nu+1}, y, w_\nu)$$
$$= h^2\frac{\partial^2 S}{\partial x^2} + h(\Delta w_\nu + \Delta w_{\nu+1})\frac{\partial^2 S}{\partial x\,\partial z} + \Delta w_\nu\,\Delta w_{\nu+1}\frac{\partial^2 S}{\partial z^2} + \frac{\partial S}{\partial z}\Delta^2 w_\nu \quad [7]$$

---

[7]) Die Argumente der zweiten Ableitungen von S sind:

$$x_{\nu+1} + (\vartheta_1 + \vartheta_2) h,\; y,\; w_\nu + \vartheta_1 \Delta w_\nu + \vartheta_2 \Delta w_{\nu+1} \qquad (0 < \vartheta_i < 1);$$

die von $\frac{\partial S}{\partial z}$:

$$x_{\nu+2},\; y,\; w_{\nu+1} + \vartheta_3(\Delta w_{\nu+1} - \Delta w_\nu) = w_{\nu+1} + \vartheta_3 \Delta^2 w_\nu.$$

ist, so folgt auf Grund von (25), daß der Betrag des in der geschweiften Klammer der Gleichung (44) stehenden Ausdrucks kleiner ist als

$$\mathrm{Max}\,|\Delta^2 w_\nu|\Big[1+\frac{h}{m}\Big(2\,\mathrm{Max}\Big|\frac{\partial R}{\partial x}\Big|+\mathrm{Max}\Big|\frac{\partial S}{\partial z}\Big|\Big)\Big]$$
$$+\frac{h^3}{m}\Big(L\,\mathrm{Max}\Big|\frac{\partial^2 R}{\partial x^2}\Big|+\mathrm{Max}\Big|\frac{\partial^2 S}{\partial x^2}\Big|+2\,L\,\mathrm{Max}\Big|\frac{\partial^2 S}{\partial x\,\partial z}\Big|+L^2\,\mathrm{Max}\Big|\frac{\partial^2 S}{\partial z^2}\Big|\Big).$$

Daher folgt aus (44), (45) unter Beachtung von (40), (41) nach Hilfssatz 1

$$|\Delta^2 w_{\nu+1}| \leqq \mathrm{Max}\,|\Delta^2 w_\nu|\,(1+K_1 h)+K_1 h^3.$$

Also ist nach Induktionsannahme (42)

$$\begin{aligned}|\Delta^2 w_{\nu+1}| &< K_1 h^2 (1+K_1 h)^{\nu+1}(1+h\nu)+h^3 K_1\\ &< K_1 h^2 (1+K_1 h)^{\nu+1}(1+h\nu)+h^2 K_1 (1+K_1 h)^{\nu+1}\cdot h\\ &= K_1 h^2 (1+K_1 h)^{\nu+1}(1+h(\nu+1)),\end{aligned}$$

womit die Behauptung (42), also auch (38) bewiesen ist.

Um nun die Existenz einer Lösung von (1), (2) zu beweisen, lassen wir h die Folge

$$h^{(n)} = X:2^n \qquad (n=1,2,\ldots)$$

durchlaufen. Die Lösung von (1′), (2′) für $h=h^{(n)}$ sei $w_\nu^{(n)}$. Ferner sei $\nu h^{(n)} = x_\nu^{(n)}$. Betrachten wir irgendeinen festen Punkt x des Intervalles $0, a$, dessen Abszisse von der Form $x=\frac{Xq}{2^n}$ $(q=0,1,\ldots;\ n=1,2,\ldots)$ ist. Ist dann

$$x = h^{(n)}\nu_n,$$

so wollen wir zunächst zeigen, daß

$$\lim_{n\to\infty} w_{\nu_n}^{(n)}(y)$$

existiert und daß die Konvergenz gleichmäßig in x und y ist. Zum Beweise wollen wir $w_{\nu_{n+1}}^{(n+1)} - w_{\nu_n}^{(n)}$ abschätzen. Setzen wir für den Augenblick der Einfachheit halber $\nu_n = p+1$, also $\nu_{n+1} = 2p+2$, so ist nach (1′), (2′)

$$(46)\qquad \frac{d^2 w_{p+1}^{(n)}}{dy^2} = R(x,y)\frac{w_{p+1}^{(n)}-w_p^{(n)}}{h^{(n)}} + S(x,y,w_p^{(n)});\quad w_{p+1}^{(n)}(0)=w_{p+1}^{(n)}(1)=0,$$

$$(47)\qquad \begin{gathered}\frac{d^2 w_{2p+2}^{(n+1)}}{dy^2} = R(x,y)\frac{w_{2p+2}^{(n+1)}-w_{2p+1}^{(n+1)}}{h^{(n+1)}} + S(x,y,w_{2p+1}^{(n+1)});\\ w_{2p+2}^{(n+1)}(0)=w_{2p+2}^{(n+1)}(1)=0.\end{gathered}$$

Wegen $h^{(n+1)} = \frac{1}{2}h^{(n)}$ ist

$$\frac{w_{2p+2}^{(n+1)}-w_{2p+1}^{(n+1)}}{h^{(n+1)}} = \frac{1}{2}\,\frac{w_{2p+2}^{(n+1)}-2\,w_{2p+1}^{(n+1)}+w_{2p}^{(n+1)}}{h^{(n+1)}} + \frac{w_{2p+2}^{(n+1)}-w_{2p}^{(n+1)}}{h^{(n)}}$$

Daher folgt aus (46), (47) durch Subtraktion, wenn noch $w_{p+1}^{(n)} - w_{2p+2}^{(n+1)} = v_{p+1}$ gesetzt wird,

(48)
$$\frac{d^2 v_{p+1}}{dy^2} - \frac{1}{h^{(n)}} R(x, y)\, v_{p+1}$$
$$= -\frac{R(x, y)}{h^{(n)}} v_p + S(x, y, w_p^{(n)}) - S(x, y, w_{2p+1}^{(n+1)}) - \frac{1}{2}\,\frac{\Delta^2 w_{2p}^{(n+1)}}{h^{(n+1)}}.$$

(49)
$$v_{p+1}(0) = v_{p+1}(1) = 0.$$

Die rechte Seite der Gleichung ist

$$-\frac{R(x, y)}{h^{(n)}}\left[v_p + \frac{h^{(n)}}{R(x, y)}\,\frac{\partial S}{\partial z}\,(w_{2p+1}^{(n+1)} - w_p^{(n)}) + \frac{1}{R(x, y)}\,\Delta^2 w_{2p}^{(n+1)}\right]\ ^{8)}$$

$$= -\frac{R(x, y)}{h^{(n)}}\left\{v_p\left[1 - \frac{h^{(n)}}{R(x, y)}\,\frac{\partial S}{\partial z}\right] + \frac{h^{(n)}\,\frac{\partial S}{\partial z}\,(w_{2p+1}^{(n+1)} - w_{2p}^{(n+1)}) + \Delta^2 w_{2p}^{(n+1)}}{R(x, y)}\right\}.$$

Da nach (25) und (38)

$$|w_{2p+1}^{(n+1)} - w_{2p}^{(n+1)}| < L h^{(n+1)} = \frac{1}{2} L h^{(n)};$$

$$|\Delta^2 w_{2p}^{(n+1)}| < L_2 (h^{(n+1)})^2 = \frac{1}{4} L_2 (h^{(n)})^2$$

ist, wird der in der geschweiften Klammer stehende Ausdruck dem Betrage nach kleiner als

$$|v_p|\left(1 + \frac{h^{(n)}}{m}\,\mathrm{Max}\left|\frac{\partial S}{\partial z}\right|\right) + \frac{1}{2m}(h^{(n)})^2\left[L\,\mathrm{Max}\left|\frac{\partial S}{\partial z}\right| + \frac{L_2}{2}\right].$$

Daher folgt aus (48), (49) nach Hilfssatz 1

(50)
$$|v_{p+1}| \leqq \mathrm{Max}\,|v_p|\,(1 + K_2 h^{(n)}) + K_2 (h^{(n)})^2,$$

sobald die von x, y, n unabhängige Konstante K_2 die Ungleichungen

(51)
$$\frac{1}{m}\,\mathrm{Max}\left|\frac{\partial S}{\partial z}\right| < K_2, \qquad \frac{1}{2m}\left[L\,\mathrm{Max}\left|\frac{\partial S}{\partial z}\right| + \frac{1}{2} L_2\right] < K_2$$

erfüllt. Da $v_0 = 0$ ist, ergibt sich aus (50) rekursiv

$$|v_{p+1}| \leqq K_2 (h^{(n)})^2 (p+1)(1 + K_2 h^{(n)})^p \leqq K_2 h^{(n)} x e^{K_2 x}.$$

Somit ist bewiesen

$$|w_{\nu_{n+1}}^{(n+1)} - w_{\nu_n}^{(n)}| \leqq h^{(n)} K_2 x e^{K_2 x} \qquad (0 \leqq x < a;\ 0 \leqq y \leqq 1).$$

Ebenso ist

$$|w_{\nu_{n+2}}^{(n+2)} - w_{\nu_{n+1}}^{(n+1)}| \leqq h^{(n+1)} K_2 x e^{K_2 x},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$|w_{\nu_{n+i}}^{(n+i)} - w_{\nu_{n+i-1}}^{(n+i-1)}| \leqq h^{(n+i-1)} K_2 x e^{K_2 x} \qquad (i = 1, 2, 3, \ldots).$$

---

[8]) Das Argument z von $\frac{\partial S}{\partial z}$ ist ein Mittelwert zwischen $w_{2p+1}^{(n+1)}$ und $w_p^{(n)}$.

Also, da $h^{(n+\mu)} = h^{(n)} : 2^\mu$,

$$(52)\qquad |w_{r_{n+i}}^{(n+i)} - w_{r_n}^{(n)}| < h^{(n)} K_2 x e^{K_1 x} \sum_{\mu=0}^{\infty} \frac{1}{2^\mu} = h^{(n)} 2 K_2 x e^{K_1 x},$$

oder erst recht

$$(53)\qquad |w_{r_{n+i}}^{(n+i)} - w_{r_n}^{(n)}| < h^{(n)} 2 K_2 X e^{K_1 X}.$$

Hieraus folgt die behauptete gleichmäßige Konvergenz der $w_{r_n}^{(n)}$. Die Grenzfunktion sei $w(x, y)$; sie ist definiert für alle y im Intervall $0 \leqq y \leqq 1$ und alle x der Form $Xq : 2^n$ im Intervall 0 bis a. Nach (25) genügen die Funktionen $w_{r_n}^{(n)}(y)$ in bezug auf x einer von y und n unabhängigen Lipschitzbedingung, so daß für die Grenzfunktion $w(x, y)$ offenbar das gleiche gilt. Hieraus wiederum folgt, daß $w(x, y)$ zu einer eindeutig bestimmten, für alle x des Intervalles $0 \leqq x \leqq a$ definierten Funktion $z(x, y)$ ergänzt werden kann, welche ihrerseits einer von y unabhängigen Lipschitzbedingung in bezug auf x genügt. Ferner folgt unter Benutzung der für jede Funktion $f(x)$ gültigen Identität

$$\begin{aligned} & f(x + 2h^{(n)}) - 2f(x + h^{(n)}) + f(x) \\ = & \sum_{\nu=0}^{2^i-1} \sum_{\mu=0}^{2^i-1} [f(x + (\nu + \mu + 2) h^{(n+i)}) - 2f(x + (\nu + \mu + 1) h^{(n+i)}) \\ & + f(x + (\nu + \mu) h^{(n+i)})] \end{aligned}$$

aus (38) leicht die Ungleichung

$$|\Delta^2 z| = |z(x + 2h, y) - 2z(x + h, y) + z(x, y)| \leqq L_2 h^2.$$

Hieraus wiederum folgt die Existenz und Stetigkeit von $\frac{\partial z}{\partial x}$. *)

---

*) Sei in der Tat $f(x)$ eine im Intervall $0, a$ stetige Funktion, für welche

$$|f(x+2h^{(n)}) - 2f(x+h^{(n)}) + f(x)| < L_2 h^{(n)2} \qquad (h^{(n)} = X : 2^n)$$

gilt; dann ist, wie wir behaupten, $f(x)$ stetig differenzierbar. Ist nämlich

$$\Delta(f, h) = f(x+h) - f(x);$$

$$\Delta^2(f, h) = f(x+2h) - 2f(x+h) + f(x),$$

so gilt identisch

$$\Delta(f, h) = \Delta^2\left(f, \frac{h}{2}\right) + 2\Delta\left(f, \frac{h}{2}\right).$$

Hieraus folgt durch wiederholte Anwendung

$$\Delta(f, h) = \sum_{\lambda=1}^{k} 2^{\lambda-1} \Delta^2\left(f, \frac{h}{2^\lambda}\right) + 2^k \Delta\left(f, \frac{h}{2^k}\right) \qquad (k = 1, 2, \ldots).$$

Setzen wir hierin $h = h^{(n)} = X : 2^n$, so folgt nach Division durch $h^{(n)}$ wegen $h^{(n)} = 2^k h^{(n+k)}$)

$$\frac{\Delta(f, h^{(n)})}{h^{(n)}} - \frac{\Delta(f, h^{(n+k)})}{h^{(n+k)}} = \frac{1}{h^{(n)}} \sum_{\lambda=1}^{k} 2^{\lambda-1} \Delta^2\left(f, \frac{h^{(n)}}{2^\lambda}\right).$$

(Fortsetzung der Fußnote *) auf nächster Seite.)

Um nunmehr zu zeigen, daß z der Differentialgleichung (1) genügt, behaupten wir zunächst: Für jedes x $(0 \leqq x \leqq a)$ der Form $Xq:2^p$ $(q = 0, 1, 2, \ldots; \; p = 1, 2, \ldots)$ ist

$$\frac{\partial z}{\partial x} = \lim_{n \to \infty} \frac{w_{\nu_n+1}^{(n)} - w_{\nu_n}^{(n)}}{h^{(n)}} \qquad (\nu_n h^{(n)} = x), \tag{54}$$

und zwar ist die Konvergenz im Bereiche (49) gleichmäßig in x und y.

Beweis. Sei $\varepsilon > 0$ eine vorgegebene Zahl. Wegen der Stetigkeit von $\frac{\partial z}{\partial x}$ können wir eine von x und y unabhängige Zahl $\bar{n}$ so wählen, daß

$$\left| \frac{\partial z(x,y)}{\partial x} - \frac{z(x + h^{(\bar{n})}, y) - z(x,y)}{h^{(\bar{n})}} \right| < \frac{\varepsilon}{3} \tag{55}$$

ist. Wir behaupten aber weiter: Wir können $\bar{n}$ auch so groß wählen, daß für $n \geqq \bar{n}$

$$\left| \frac{w_{\bar{\nu}_n}^{(n)} - w_{\nu_n}^{(n)}}{h^{(\bar{n})}} - \frac{w_{\nu_n+1}^{(n)} - w_{\nu_n}^{(n)}}{h^{(n)}} \right| < \frac{\varepsilon}{3} \qquad \left(\begin{aligned} \bar{\nu}_n h^{(n)} &= x + h^{(\bar{n})} \\ \nu_n h^{(n)} &= x \\ (\nu_n + 1) h^{(n)} &= x + h^{(n)} \end{aligned} \right) \tag{56}$$

denn die gleiche Betrachtung wie die in Anm. [9]) zeigt auf Grund von (38), daß die links stehende Differenz kleiner als $h^{(\bar{n})} L_2 : 2$ (vgl. die Ungleichung ** der Anm.) ist. Wegen der gleichmäßigen Konvergenz können wir ferner zu $h^{(\bar{n})}$ ein $n_1 \geqq \bar{n}$ so wählen, daß für $n \geqq n_1$

$$\left| \frac{z(x + h^{(\bar{n})}, y) - z(x,y)}{h^{(\bar{n})}} - \frac{w_{\bar{\nu}_n}^{(n)} - w_{\nu}^{(n)}}{h^{(\bar{n})}} \right| < \frac{\varepsilon}{3} \qquad (\bar{\nu}_n h^{(n)} = x + h^{(\bar{n})}). \tag{57}$$

Addition der Ungleichungen (55) bis (57) liefert aber die Behauptung (54).

Also ist nach *

$$** \qquad \left| \frac{\Delta(f, h^{(n)})}{h^{(n)}} - \frac{\Delta(f, h^{(n+k)})}{h^{(n+k)}} \right| < \frac{L_2}{h^{(n)}} \sum_{\lambda=1}^{k} \frac{2^{\lambda-1} (h^{(n)})^2}{2^{2\lambda}} = \frac{L_2 h^{(n)}}{2} \sum_{\lambda=1}^{k} \frac{1}{2^\lambda} < h^{(n)} \frac{L_2}{2}.$$

Aus dieser Ungleichung folgt aber nach dem Cauchyschen Konvergenzprinzip, daß die Folge der stetigen Funktionen von x

$$f_n(x) = \frac{\Delta(f, h^{(n)})}{h^{(n)}} = \frac{f(x + h^{(n)}) - f(x)}{h^{(n)}}$$

gleichmäßig konvergiert. Die Grenzfunktion, die nach Definition eine spezielle rechte Derivierte von $f(x)$ ist, ist also stetig. Aus der Existenz *einer* stetigen Derivierten folgt aber nach einem Satz aus der Theorie der reellen Funktionen, daß $f(x)$ stetig differenzierbar ist (s. Carathéodory, Vorlesungen über reelle Funktionen, 1. Aufl., S. 534, Satz 7). In unserem Falle ist $f(x) = z(x, y)$; da die Abschätzung *, also auch ** unabhängig von y gilt, ist die Konvergenz der f_n gleichmäßig in x *und* y, also $\frac{\partial z}{\partial x}$ stetig in (x, y).

Für jedes positive $x < a$ der Form $Xq:2^p$ besteht nun von einem gewissen n ab die Gleichung

$$\frac{d^2 w^{(n)}_{\nu_n+1}}{dy^2} = R(x+h^{(n)}, y)\,\frac{w^{(n)}_{\nu_n+1} - w^{(n)}_{\nu_n}}{h^{(n)}} + S(x+h^{(n)}, y, w^{(n)}_{\nu_n}).$$

Nach dem Bewiesenen existiert der Limes der rechten Seite für $n = \infty$, und zwar ist die Konvergenz gleichmäßig in (x, y). Daher konvergiert auch $\frac{d^2 w^{(n)}_{\nu_n+1}}{dy^2}$ gleichmäßig gegen eine stetige Grenzfunktion $W(x, y)$, die für alle x der Form $Xq:2^p$ erklärt ist, und es besteht nach (54) für alle x dieser Form die Gleichung

$$W(x, y) = R(x, y)\,\frac{\partial z}{\partial x} + S(x, y, z(x, y)).$$

Die rechte Seite ist für alle positiven $x < a$ definiert und stetig, und wir definieren $W(x, y)$ für alle diese x durch diese Gleichung. Es ist nun zu zeigen, daß $\frac{\partial^2 z}{\partial y^2}$ existiert und gleich $W(x, y)$ ist. Für alle x der Form $Xq:2^p$ ist das klar[10]). Ist x ein beliebiger Wert des Intervalles $0, a$, so sei x_k eine Folge gegen x konvergierender Zahlen der Form $Xq:2^q$. Dann ist auf Grund des schon Bewiesenen gleichmäßig in y.

$$\lim_{k\to\infty} z(x_k, y) = z(x, y),$$

$$\lim_{k\to\infty} \frac{\partial^2 z(x_k, y)}{\partial y^2} = W(x, y).$$

Hieraus folgt die Behauptung (vgl. Anm. [10])).

Hiermit ist gezeigt, daß die Funktion z die Gleichung (1) erfüllt. Daß die Randbedingungen (2) erfüllt sind, ist klar, da sie von allen $w^{(n)}_{\nu_n}$ erfüllt sind und z stetig ist. — Satz 2 ist somit bewiesen.

Bemerkung über den Fehler $z(x_\nu, y) - w_\nu(y)$. Eine Abschätzung dieses Fehlers liegt schon in Gleichung (22), S. 655, vor. Jedoch hängt

---

[10]) Ist gleichmäßig in $0 \leq y \leq 1$ $\lim\limits_{n\to\infty} g_n(y) = g(y)$, ist ferner g_n zweimal stetig differenzierbar und ist gleichmäßig $\lim\limits_{n\to\infty} g_n'' = G(y)$, so existiert $g''(y)$ und es ist $G(y) = g''(y)$. Denn es ist

$$g_n(y) = a_n + b_n y + \int_0^y\!\!\int_0^\eta g_n''(\zeta)\,d\zeta.$$

Hieraus folgt auf Grund der Voraussetzungen zunächst, daß $\lim\limits_{n\to\infty} a_n + b_n y$ existiert also auch $\lim\limits_{n\to\infty} a_n = a$, $\lim\limits_{n\to\infty} b_n' = b$ existieren. Also wird

$$g(y) = a + by + \int_0^y\!\!\int_0^\eta G(\zeta)\,d\zeta,$$

woraus die Behauptung folgt.

die dort auftretende Zahl τ von der durch (14) definierten Funktion $\eta_\nu(y)$ ab. Die Benutzung dieser Abschätzung würde daher schon die Kenntnis der gesuchten Lösung z voraussetzen. Dagegen erhalten wir aus (52) die Abschätzung

$$|z(x,y) - w_{\nu_n}^{(n)}(y)| < h^{(n)} 2 K_2 x e^{K_1 x} \qquad (h^{(n)} \nu_n = x),$$

wo K_2, wie aus dem Beweis für Satz 2 hervorgeht, aus den gegebenen Funktionen R, S, z_0 und deren Ableitungen berechenbar ist: Man bestimme zuerst m (Gleichung (10)), dann K (Gleichung (29) bis (31)), dann L (Gleichung (34)), dann L_1 (S. 659), dann K_1 (Gleichung (39) bis (41)), dann L_2 (S. 659) und schließlich K_2 nach (51).

IV.

Satz 3. *Vorgelegt sei das Problem* (1), (2). *Die Koeffizienten von* (1) *mögen den in Satz* 2 *gemachten Voraussetzungen genügen. Außerdem werde noch Existenz und Stetigkeit von* $\frac{\partial^2 S}{\partial y \partial z}$ *sowie der Ableitungen von* $R(x,y)$ *bis zu den vierten vorausgesetzt. Über* $z_0(y)$ *werde dagegen außer* (23) *nur Stetigkeit vorausgesetzt. Ist dann* a *eine den Ungleichungen* (27), (28) *genügende positive Konstante, so existiert eine der Randbedingung* (2) *genügende Funktion* $z(x,y)$, *die in jedem inneren Punkte von* (11) *eine Lösung von* (1) *ist.*

Beweis. Wir wollen zunächst eine Folge von Funktionen $z_0^{(n)}(y)$ herstellen, die gleichmäßig in $0 \leq y \leq 1$ gegen $z_0(y)$ konvergieren und von denen jede die in Satz 2 über z_0 gemachten Voraussetzungen erfüllt. Sei $s(y)$ eine zweimal stetig differenzierbare Funktion, die den Bedingungen

$$s(0) = S(0,0,0); \quad s(1) = S(0,1,0) \tag{58}$$

genügt. Die Funktion

$$\begin{gathered} F(y) = z_0(y) + \int_0^1 s(\eta)\gamma(y,\eta)\,d\eta \\ \left(\gamma(y,\eta) = \begin{cases} (1-\eta)y & \text{für } y \leq \eta \\ \eta(1-y) & \quad\text{„}\quad y \geq \eta \end{cases}\right) \end{gathered} \tag{59}$$

ist dann jedenfalls in $0 \leq y \leq 1$ stetig und verschwindet für $y = 0$ und $y = 1$. Erweitern wir den Definitionsbereich durch die Festsetzung $F(-y) = -F(y)$, $F(y+2) = F(y)$, so erhalten wir eine überall stetige ungerade periodische Funktion. Die Fejérschen aus den Fourierkoeffizienten von $F(y)$ gebildeten Mittel $A_n(y)$ enthalten daher nur Sinusglieder und verschwinden nebst allen Ableitungen *gerader* Ordnung für $y = 0$, $y = 1$; außerdem ist gleichmäßig in $(0, 1)$

$$F(y) = \lim_{n \to \infty} A_n(y). \tag{60}$$

Wir setzen nun

(61) $$z_0^{(n)}(y) = A_n(y) - \int_0^1 s(\eta)\gamma(y, \eta)\, d\eta.$$

Dann ist

(62) $$z_0^{(n)}(0) = z_0^{(n)}(1) = 0$$

und

(63) $$z_0^{(n)\prime\prime}(y) = A_n''(y) + s(y),$$

also

$$z_0^{(n)\prime\prime} = s(0); \qquad z^{(n)\prime\prime}(1) = s(1)$$

oder nach (58) und (62)

$$z_0^{(n)\prime\prime}(0) = S(0, 0, z_0^{(n)}(0)), \qquad z_0^{(n)\prime\prime}(1) = S(0, 1, z_0^{(n)}(1)).$$

Ferner ist $z_0^{(n)}(y)$ nach (63) viermal stetig differenzierbar, da A_n'' und s zweimal stetig differenzierbar sind. Da außerdem nach (59) bis (61) gleichmäßig

(64) $$\lim_{n \to \infty} z_0^{(n)}(y) = z_0(y)$$

ist, ist für genügend großes n nicht nur nach (23)

$$|z_0^{(n)}| < Z,$$

sondern es gilt auch die Ungleichung (28) noch, wenn in ihr z_0 durch $z_0^{(n)}$ ersetzt wird. Im folgenden soll n immer so groß gewählt sein. Nach dem Vorstehenden erfüllt dann $z_0^{(n)}(y)$ alle in Satz 2 von z_0 verlangten Voraussetzungen. Nach diesem Satz existiert daher im abgeschlossenen Bereiche (11) eine Lösung $z^{(n)}(x, y)$ von (1), die die Randbedingungen

(65) $$z^{(n)}(0, y) = z_0^{(n)}(y); \qquad z^{(n)}(x, 0) = z^{(n)}(x, 1) = 0$$

genügt. Können wir nun zeigen, daß die $z^{(n)}(x, y)$ in dem abgeschlossenen Bereiche (11) gleichmäßig konvergieren und daß $\frac{\partial z^{(n)}}{\partial y}$ und $\frac{\partial^2 z^{(n)}}{\partial y^2}$ in jedem abgeschlossenen Teilbereiche von (11) gleichmäßig konvergieren, so genügt $z(x, y) = \lim_{n \to \infty} z^{(n)}(x, y)$ offenbar allen Forderungen unseres Satzes.

Für den Nachweis der Konvergenz ist die Annahme

(66) $$R(x, y) \equiv 1$$

keine Beschränkung der Allgemeinheit, wie man leicht ersieht, wenn man die Transformation

$$X = \int_0^x \frac{d\xi}{\left(\int_0^1 \sqrt{R(\xi, \eta)}\, d\eta\right)^2}, \qquad Y = \frac{1}{\int_0^1 \sqrt{R(x, \eta)}\, d\eta} \int_0^y \sqrt{R(x, \eta)}\, d\eta \text{ [11]}$$

[11]) Block, Sur les équations aux dérivées partielles du type parabolique, Arkiv för Matematik 6 (1911), Nr. 31, S. 4.

und die weitere

$$\zeta = z\, e^{-\frac{1}{2}\int\limits_0^Y T_1(X,\eta)\,d\eta}$$

ausführt, wo T_1 der Koeffizient von $\frac{\partial z}{\partial Y}$ in der durch die erste Transformation entstandenen Gleichung ist. Im folgenden soll daher (66) angenommen werden.

Wir setzen nun

(67) $$z^{(n+p)}(x, y) - z^{(n)}(x, y) = v^{(n,p)}(x, y).$$

Dann genügt $v^{(n,p)}$ der Gleichung

(68) $$\frac{\partial^2 v}{\partial y^2} = \frac{\partial v}{\partial x} + v \cdot \sigma^{(n,p)}(x, y),$$

wo

(69) $$|\sigma^{(n,p)}| = \left|\frac{\partial S}{\partial z}(x, y, \tilde{z}(x, y))\right| \leqq \operatorname{Max}\left|\frac{\partial S}{\partial z}\right| = M$$

($\tilde{z}$ Zwischenwert zwischen $z^{(n)}$ und $z^{(n+p)}$),

und den Randbedingungen

(70) $$v(0, y) = z_0^{(n+p)}(y) - z_0^{(n)}(y), \qquad v(0, x) = v(1, x) = 0.$$

Da wir wissen, daß (68), (70) die Lösung $v = v^{(n,p)}$ hat, können wir diese nach Satz 1 durch Grenzübergang aus

(71) $$\frac{d^2 v_{\nu+1}}{dy^2} - \frac{1}{h} v_{\nu+1} = -\frac{v_\nu}{h} + v_\nu \sigma^{(n,p)}(x_{\nu+1}, y)$$

(72) $$v_\nu(0) = v_{\nu+1}(1) = 0, \qquad v_0(y) = z_0^{(n+p)}(y) - z_0^{(n)}(y)$$

erhalten. Nach Hilfssatz 1 ist nun

$$|v_{\nu+1}| \leqq \operatorname{Max}|v_\nu(1 + h\sigma^{(n,p)})|,$$

also nach (69)

(73) $$\operatorname{Max}|v_{\nu+1}| \leqq \operatorname{Max}|v_\nu|(1 + M h).$$

Wegen (64) können wir ferner nach der letzten Gleichung (72)

$$|v_0(y)| < \varepsilon_n$$

mit

(74) $$\lim_{n \to \infty} \varepsilon_n = 0$$

schreiben, so daß aus (73) durch vollständige Induktion $\operatorname{Max}|v_{\nu+1}| \leqq \varepsilon_n(1 + Mh)^\nu$ folgt. Hieraus wieder ergibt sich (ähnlich wie beim Beweis von 22)

$$|v_{\nu+1}| \leqq \varepsilon_n e^{xM}.$$

Daher ist nach (67)

(75) $$|v^{(n,p)}| = |z^{(n+p)}(x, y) - z^{(n)}(x, y)| \leqq \varepsilon_n e^{xM},$$

woraus nach (74) die behauptete in (11) gleichmäßige Konvergenz von $z^{(n)}$ folgt. Hieraus folgt auch, daß die stetige Grenzfunktion $z(x, y)$ die Randbedingungen erfüllt. Denn es ist

$$\lim_{n\to\infty} z^{(n)}(0, y) = z_0(y), \quad \lim_{n\to\infty} z^{(n)}(x, 0) = z^{(n)}(x, 0) = 0,$$
$$\lim_{n\to\infty} z^{(n)}(x, 1) = z^{n}(x, 1) = 0.$$

Um nun weiter zu zeigen, daß $\frac{\partial z^{(n)}}{\partial y}$, $\frac{\partial^2 z^{(n)}}{\partial y^2}$ in jedem abgeschlossenen Teilbereich von (11) gleichmäßig konvergieren, bedienen wir uns der zum Bereich (11) gehörenden Greenschen Funktion $G(x, y, \xi, \eta)$ von $\frac{\partial^2 z}{\partial y^2} - \frac{\partial z}{\partial x}$, die bekanntlich folgendermaßen definiert ist: Ist

$$U(x, y, \xi, \eta) = \begin{cases} \frac{1}{\sqrt{x-\xi}} e^{-\frac{(y-\eta)^2}{4(x-\xi)}} & \text{für } \xi < x, \\ 0 & \text{„} \quad \xi \geqq x, \end{cases}$$

so ist

$$G(x, y, \xi, \eta) = U(x, y, \xi, \eta) - g(x, y, \xi, \eta),$$

wobei g (als Funktion von ξ, η) der Differentialgleichung

$$\frac{\partial^2 g}{\partial \eta^2} + \frac{\partial g}{\partial \xi} = 0$$

und den Randbedingungen

$$g(x, y, x, \eta) = 0; \quad g(x y, \xi, 0) = U(x, y, \xi, 0); \quad g(x, y, \xi, 1) = U(x, y, \xi, 1)$$

genügt. Die *Existenz* von g, und damit auch von G, folgt (für jedes im Innern von (11) gelegene x, y) leicht aus Satz 2 [12]).

$z^{(n)}$ ist nun nach Satz 2 auch am Rande von (11) noch differenzierbar. Man erhält daher durch bekannte Anwendung einer Greenschen Formel (unter Berücksichtigung von (65))

$$(76) \qquad 2\sqrt{\pi}\, z^{(n)}(x, y) = \int_0^1 G(x, y, 0, \eta)\, z_0^{(n)}(\eta)\, d\eta$$
$$- \int_0^x \int_0^1 G(x, y, \xi, \eta)\, S(\xi, \eta, z^{(n)}(\xi, \eta))\, d\xi\, d\eta$$

und nach (67)

$$(77) \qquad 2\sqrt{\pi}\, v^{(n,p)}(x, y) = \int_0^1 G(x, y, 0, \eta)\, v^{(n,p)}(0, \eta)\, d\eta$$
$$- \int_0^x \int_0^1 G(x, y, \xi, \eta) [S(\xi, \eta, z^{(n+p)}) - S(\xi, \eta, z^{(n)})]\, d\xi\, d\eta.$$

---

[12]) Setzt man nämlich $\xi' = x - \xi$, $\eta' = \eta$ und

$$g(x, y, \xi, \eta) - (1-\eta)\, U(x, y, \xi, 0) - \eta\, U(x, y, \xi, 1) = \gamma(x, y, \xi', \eta'),$$

so erhält man, wie die Rechnung zeigt, für γ als Funktion von ξ', η' ein allen Voraussetzungen von Satz 2 genügendes Problem. — Der weitere Beweis des Textes bedient sich bekannter Methoden und ist daher nur in den Hauptzügen kurz wiedergegeben.

Auf Grund bekannter Abschätzungen[13]) der Integrale

$$\int_0^1 \frac{\partial U}{\partial y}(x, y, 0, \eta)\, d\eta, \qquad \int_0^x\int_0^1 \frac{\partial U}{\partial y}(x, y, \xi, \eta)\, f(\xi, \eta)\, d\xi\, d\eta$$

ergeben sich nun aus (76), (77) unter Benutzung von (75) in einem abgeschlossenen Teilbereich $\mathfrak{T}$ von (11) ohne Mühe Ungleichungen der Form

$$(78) \qquad \left|\frac{\partial z^{(n)}}{\partial y}\right| < \frac{M}{\sqrt{x}}, \qquad \left|\frac{\partial v^{(n,p)}}{\partial y}\right| < \frac{\eta_n}{\sqrt{x}},$$

wo M eine von n unabhängige Konstante und η_n eine Folge gegen 0 konvergierender (von p unabhängiger) Konstanten ist. Unter Benutzung der gleichen Abschätzungen und von (78), sowie der Tatsache, daß der Faktor von G in dem Doppelintegral von (77) nach η differenzierbar ist, findet man schließlich leicht, daß auch $\frac{\partial^2 v^{(n,p)}}{\partial y^2}$ in $\mathfrak{T}$ gleichmäßig für $n \to \infty$ gegen 0 konvergiert. Hiermit ist die gleichmäßige Konvergenz von $\frac{\partial z^{(n)}}{\partial y}$ und $\frac{\partial^2 z^{(n)}}{\partial y^2}$ in $\mathfrak{T}$, und somit nach Satz 3, S. 666f. bewiesen.

---

[13]) Siehe z. B. Gevrey, loc. cit. S. 330 und S. 343.

(Eingegangen am 5. 5. 1929.)

Über das Dämpfungsproblem der mathematischen Physik, mit einer Anwendung auf die Akustik großer Räume[1]).

Von

M. J. O. Strutt in Eindhoven (Niederlande).

Durch die Arbeiten von Poincaré, Fredholm, Hilbert u. a. sind die Entwicklungssätze für Eigenwertaufgaben mit reellen, symmetrischen Kernen, also *reellen* Eigenwerten im weiten Umfange sichergestellt. Die anschließende und in der Physik für die Debijesche Theorie der spezifischen Wärme und die Jeanssche Theorie der Hohlraumstrahlung wichtige Aufgabe der asymptotischen Berechnung von Eigenwerten für beliebige Gebiete wurde von Weyl mittels der Theorie der Integralgleichungen und von Courant durch Anwendung der Variationsrechnung gelöst. Auch für akustische Probleme ist diese asymptotische Kenntnis der Eigenwerte, wie Sommerfeld dartat, prinzipiell wichtig.

In Wirklichkeit sind die in der mathematischen Physik auftretenden Schwingungsaufgaben vielfach mit Dämpfung behaftet, was sich stets in einem Ausklingen angeregter freier Schwingungen kundtut.

Die namentlich für die Akustik wichtige Aufgabe der Berechnung dieser An- und Abklingungsvorgänge, sowie der unter Zwang sich einstellenden stationären Amplituden führt auf Entwicklungsprobleme, die über den durch die oben erwähnten Arbeiten behandelten Bereich hinausgreifen.

W. Sabine [2]) hat in der experimentellen Akustik einen Satz entdeckt, der für die Konstruktion von Räumen von grundlegender Bedeutung ist. Er lautet, daß die *Nachhalldauer eines genügend großen Raumes lediglich von der gesamten Oberfläche des dämpfenden Materiales und vom Volumen abhängt, nicht aber von der Form.* Es liegt nahe, diesen Satz dem von Lorentz ausgesprochenen und von Weyl bewiesenen an die Seite zu stellen,

---

[1]) Mathematische Bearbeitung eines Vortrages, der in der «Nederlandsche Natuurkundige Vereeniging» zu Amsterdam am 26. Januar 1929 gehalten wurde.

[2]) Collected Papers, S. 34.

nachdem asymptotisch die reellen Eigenwerte eines Raumes nur vom Volumen abhängen. Der Sabinesche Satz, der bisher nur durch Reflexionsbetrachtungen, denen unübersehbare Phasenverhältnisse anhaften, bewiesen wurde, stellt sich als eine asymptotische Eigenschaft der *komplexen* Eigenwerte *gedämpfter* Schwingungsprobleme heraus.

Im Gegensatz zur reichen Literatur über ungedämpfte Aufgaben ist jene über gedämpfte Probleme auffallend arm. O. Faber[8]) hat in einer grundlegenden Arbeit für eindimensionale Gebiete Entwicklungssätze aufgestellt und asymptotisch die Eigenwerte berechnet. Seine Beweisführung stützt sich wesentlich auf die asymptotische Kenntnis der Eigenfunktionen und der Greenschen Funktion für die eindimensionale Aufgabe. Für mehrere Dimensionen besitzen wir diese Kenntnis nicht und ist also ein vom Faberschen abweichender, allgemein gehaltener Rechnungsgang notwendig. Ziel der vorliegenden Arbeit ist es, gestützt auf eine einzige Integralgleichung, sowohl die asymptotische Berechnung der komplexen Eigenwerte, des Verhaltens der Eigenfunktionen vorzunehmen, als die Entwicklungssätze für gedämpfte Schwingungsaufgaben abzuleiten. Die Form der hier gegebenen Entwicklungen weicht von derjenigen Fabers ab. Anschließend wird als Anwendung der oben erwähnte Sabinesche Satz bewiesen.

I. Asymptotisches Verhalten der Eigenwerte und Eigenfunktionen.

Das *Dämpfungsproblem der mathematischen Physik* kann folgendermaßen formuliert werden. Gesucht ist eine Lösung der Differentialgleichung

$$(1)\quad L[u] \equiv \frac{\partial}{\partial x}\left(p(xyz)\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(p\frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial z}\left(p\frac{\partial u}{\partial z}\right) - q(xyz)\,u$$
$$= \varrho(xyz)\frac{\partial^2 u}{\partial t^2} + w(xyz)\frac{\partial u}{\partial t} \qquad [p > 0],$$

welche auf der Gebietsgrenze eines ganz im Endlichen gelegenen Gebietes der Gleichung

$$(2)\qquad \sigma(xyz)\,u + \frac{\partial u}{\partial n} = 0 \quad (n \text{ äußere Normale})$$

genügt. Über die Begrenzung des Gebietes machen wir dabei die Voraussetzung, daß sie aus endlich vielen stetigen Flächen zusammengesetzt sei. Um das Problem zu einem vollständig bestimmten zu machen sind noch zwei Anfangsbedingungen für einen fest gewählten Punkt $t = t_0$ vorzugeben. Doch stellen wir dieses Anfangswertproblem und die zu ihrer Lösung notwendigen Entwicklungssätze bis Abschnitt 3 zurück.

---

[8]) Diss. Straßburg 1914 (Ref. v. Mises). Vgl. auch A. C. Dixon, Proc. London math. Soc. (2) **3** (1905), S. 83, sowie T. J. I'A. Bromwich, ibid. (2) **15** (1914) S. 444–446.

Wir setzen:

$$u = v(xyz)\,e^{-j\sqrt{\lambda}\,t} \qquad (j = \sqrt{-1}),$$

wodurch (1) sich schreibt:

$$L(v) + \varrho\lambda v + jw\sqrt{\lambda}\,v = 0. \tag{1a}$$

Man erkennt leicht, daß diese Separation die einzig mögliche ist.

Gesucht sind also zunächst die Eigenwerte und Eigenfunktionen der homogenen Randaufgabe (1a), (2).

Unter Benutzung einer Greenschen Funktion

$$G(xyz, \xi\eta\zeta),$$

die den Bedingungen:

a) G genügt überall im Gebiet der Gleichung (1a) mit $\lambda = 0$ in xyz und in $\xi\eta\zeta$ außer im Punkte $x = \xi$, $y = \eta$, $z = \zeta$;

b) G genügt auf der Gebietsgrenze der Gleichung (2), außer für $x = \xi$, $y = \eta$, $z = \zeta$;

c) Im Punkte $x = \xi$, $y = \eta$, $z = \zeta$ verhält G sich so, daß ihre Normalableitung, integriert über einer kleinen Kugelfläche K um diesen Punkt herum die Bedingung

$$\lim_{K \to 0} \oint_K \frac{\partial G}{\partial n}\,dK = -\frac{1}{p}$$

befriedigt;

d) G ist eine stetige Funktion von xyz und $\xi\eta\zeta$

genügt, läßt sich diese Randaufgabe als eine lineare homogene Integralgleichung:

$$\begin{aligned} v(xyz) = {} & \lambda \iiint G(xyz, \xi\eta\zeta)\,\varrho(\xi\eta\zeta)\,v(\xi\eta\zeta)\,d\xi\,d\eta\,d\zeta \\ & + j\sqrt{\lambda} \iiint G(xyz, \xi\eta\zeta)\,w(\xi\eta\zeta)\,v(\xi\eta\zeta)\,d\xi\,d\eta\,d\zeta \end{aligned} \tag{3}$$

schreiben [4]).

Vorausgesetzt, daß ϱ durchweg positiv ist:

$$\varrho > 0,$$

kann man statt (3) schreiben:

$$\begin{aligned} v &= \lambda \iiint G\varrho v\left[1 + j\frac{w}{\varrho\sqrt{\lambda}}\right] d\xi\,d\eta\,d\zeta \\ &= \lambda\left\{\iiint G\varrho v\,d\xi\,d\eta\,d\zeta + O\left(\frac{1}{\sqrt{|\lambda|}}\right)\right\}. \end{aligned} \tag{3a}$$

Aus (3a) geht hervor, daß die Eigenwerte und Eigenfunktionen von (3) bis auf Glieder der Ordnung $|\lambda^{-\frac{1}{2}}|$, wie man durch Störungsrechnung be-

---

[4]) Aus (3) ist abzuleiten, daß die Eigenfunktionen v_i *nicht orthogonal* im gewöhnlichen Sinne zueinander sind.

weist, für $\lim |\lambda| \to \infty$ mit den reellen Eigenwerten und Eigenfunktionen der Gleichung

$$(4) \qquad v = \lambda \iiint G \varrho v \, d\xi \, d\eta \, d\zeta,$$

zusammenfallen.

Insbesondere gelten also asymptotisch für die Eigenwerte von (3) die von Weyl und Courant abgeleiteten Formeln.

Aus (3a) geht hervor, daß die Eigenwerte von (3) sich asymptotisch schreiben

$$(5) \qquad +\sqrt{\lambda} = \alpha + j\beta,$$

wo

$$\lim_{|\lambda| \to \infty} \left| \frac{\beta}{\alpha} \right| = O\left(\frac{1}{\alpha}\right).$$

Diese Gleichungen gestatten, asymptotisch auch den *Dämpfungsteil* β der Eigenwerte zu berechnen. Ich werde dies für eine, zwei und drei Dimensionen ausführen.

a) Eine Dimension.

Ohne Dämpfung gilt hier für die Eigenwerte von (4):

$$(6) \qquad \lim_{n \to \infty} \lambda = \alpha^2 = \frac{n^2 \pi^2}{\left(\int \sqrt{\frac{\varrho}{p}}\, dx\right)^2} + O(1) \quad (n \text{ ganze Zahl}).$$

Mit Dämpfung ist nach (3a) ϱ zu ersetzen durch

$$\varrho\left(1 + j \frac{w}{\varrho\sqrt{\lambda}}\right),$$

also $\sqrt{\varrho}$ durch

$$\sqrt{\varrho}\left(1 + j \frac{w}{2\varrho\sqrt{\lambda}}\right),$$

so daß aus (6) unter Benutzung von (5) folgt:

$$\alpha^2 + 2\alpha\beta j = \frac{n^2\pi^2}{\left(\int \sqrt{\frac{\varrho}{p}}\, dx + \frac{j}{2\alpha} \int \frac{w}{\sqrt{\varrho p}}\, dx\right)^2} = \alpha^2 - j\alpha \frac{\int \frac{w}{\sqrt{\varrho p}}\, dx}{\int \sqrt{\frac{\varrho}{p}}\, dx},$$

also:

$$(7) \qquad \lim_{\alpha \to \infty} \beta = -\frac{1}{2} \frac{\int \frac{w}{\sqrt{\varrho p}}\, dx}{\int \sqrt{\frac{\varrho}{p}}\, dx} + O\left(\frac{1}{\alpha}\right).$$

Hiermit haben wir Fabers Formeln wiedergefunden.

b) Zwei Dimensionen.

Aus den Eigenwerten von (4) ohne Dämpfung

$$(8)\qquad \lim_{|\lambda|\to\infty} \lambda = \alpha^2 = \frac{4n\pi}{\iint \frac{\varrho}{p}\,dx\,dy} + O(\alpha\ln\alpha)$$

schließt man ähnlich wie oben:

$$\alpha^2 + 2\alpha\beta j = \alpha^2 - j\alpha \frac{\iint \frac{w}{p}\,dx\,dy}{\iint \frac{\varrho}{p}\,dx\,dy},$$

also:

$$(9)\qquad \lim_{\alpha\to\infty} \beta = -\frac{1}{2}\frac{\iint \frac{w}{p}\,dx\,dy}{\iint \frac{\varrho}{p}\,dx\,dy} + O\left(\frac{\ln\alpha}{\alpha}\right).$$

c) Drei Dimensionen.

Hier liefert die Abschätzung

$$(10)\qquad \lim_{|\lambda|\to\infty} \lambda^{\frac{3}{2}} = \alpha^3 = \frac{6\pi^2 n}{\iiint \left(\frac{\varrho}{p}\right)^{\frac{3}{2}} dx\,dy\,dz} + O(\alpha^2\ln\alpha)$$

in ähnlicher Weise

$$(11)\qquad \lim_{\alpha\to\infty} \beta = -\frac{1}{2}\frac{\iiint \frac{w}{p}\sqrt{\frac{\varrho}{p}}\,dx\,dy\,dz}{\iiint \left(\frac{\varrho}{p}\right)^{\frac{3}{2}} dx\,dy\,dz} + O\left(\frac{\ln\alpha}{\alpha}\right).$$

Die Formeln (3) bis (11) lösen das in diesem Abschnitt gestellte Problem. Ihre Betrachtung erlaubt, sofort asymptotisch die komplexen Eigenwerte von (3) für beliebig viele Dimensionen hinzuschreiben.

II. Die allgemeine Greensche Funktion.

Wir machen im folgenden einen wesentlichen Gebrauch von der verallgemeinerten Greenschen Funktion (Resolvente oder Grundlösung), die zuerst von Poincaré [5]) eingeführt wurde, und definieren sie als eine Funktion

$$K(xyz, \xi\eta\zeta, \lambda),$$

welche die nachfolgenden Bedingungen erfüllt:

a) Sie genügt in xyz und in $\xi\eta\zeta$ der Gleichung (1a) im ganzen Gebiet außer für $x=\xi$, $y=\eta$, $z=\zeta$;

[5]) Rendiconti del Circolo matematico di Palermo 8 (1894), S. 189.

b) K genügt der Bedingung (2) wieder mit Ausnahme des Punktes $x=\xi$, $y=\eta$, $z=\zeta$;

c) Im Punkte $x=\xi$, $y=\eta$, $z=\zeta$ verhält K sich so, daß ihre Normalableitung, integriert über einer kleinen Kugelfläche F um diesen Punkt herum der Bedingung

$$\lim_{F\to 0}\oint_F \frac{\partial K}{\partial n}\,dF = -\frac{1}{p}$$

genügt;

d) K ist eine stetige Funktion von $x\,y\,z$ und $\xi\,\eta\,\zeta$.

Diese Greensche Funktion läßt sich mit Hilfe der früher eingeführten gewöhnlichen Greenschen Funktion G bestimmen aus der Gleichung

$$(12)\quad \begin{aligned} K(x\,y\,z,\xi\,\eta\,\zeta,\lambda) &= G(x\,y\,z,\xi\,\eta\,\zeta)+\lambda\iiint G(x\,y\,z,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,\varrho(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z}) \\ &\times K(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z},\xi\,\eta\,\zeta,\lambda)\,d\mathfrak{x}\,d\mathfrak{y}\,d\mathfrak{z}+j\sqrt{\lambda}\iiint G(x\,y\,z,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z}) \\ &\times w(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,K(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z},\xi\,\eta\,\zeta,\lambda)\,d\mathfrak{x}\,d\mathfrak{y}\,d\mathfrak{z}, \end{aligned}$$

wie man durch Operieren mit L beweist und die unter Heranziehung der Fredholmschen Theorie erlaubt, die Existenz und den meromorphen Charakter von K einzusehen, sobald man die Existenz von G als bewiesen annimmt. Analog, wie (3) gestattete, auf (3a) und (4) zu schließen, erhellt aus (12), daß K asymptotisch für $\lim|\lambda|\to\infty$ zusammenfallen wird mit der verallgemeinerten Greenschen Funktion K^*, die der Differentialgleichung

$$L[K^*]+\lambda\varrho K^* = L[G]$$

genügt und somit zu (1a) mit $w=0$ gehört, welche Funktion K^* sich, wie bekannt, in eine konvergente bilineare Reihe nach den Eigenfunktionen von (4) entwickeln läßt[6]):

$$(13)\qquad K^*(x\,y\,z,\xi\,\eta\,\zeta,\lambda)=\sum_{i=1}^{\infty}\frac{v_i(x\,y\,z)\,v_i(\xi\,\eta\,\zeta)}{(\lambda_i-\lambda)\iiint\varrho\, v_i^2\,dx\,dy\,dz}.$$

Wir schließen, daß die Entwicklung von K für sehr große $|\lambda|$ in (13) übergehen muß bis auf einen Fehler der Ordnung $|\lambda^{-\frac{1}{2}}|$.

Hiermit ist die Grundlage geschaffen für die Anwendung des Cauchyschen Residuensatzes (Partialbruchzerlegung) zur Entwicklung der Funktion K nach ihren Polen in der komplexen λ-Ebene.

Durch Betrachtung der asymptotischen Eigenwertabschätzungen des vorigen Abschnittes erhellt, daß sich in dieser Ebene von einem gewissen $|\lambda|$ an eine nach Unendlich gehende Folge von Konturen k_n angeben läßt, die keinen Eigenwert treffen.

---

[6]) Z. B. A. Sommerfeld, Phys. Zeitschr. 11 (1910), S. 1057.

Um zu beweisen, daß die Cauchysche Formel

$$\frac{1}{2\pi j}\oint\limits_{k_n}\frac{K(xyz,\xi\eta\zeta,\mu)}{\mu-\lambda}d\mu = K(xyz,\xi\eta\zeta,\lambda)+\sum_{i=1}^{n}\frac{R(\lambda_i)}{\lambda-\lambda_i},$$

wo $R(\lambda_i)$ das Residuum von K für $\lambda=\lambda_i$ bedeutet und die Summe über alle innerhalb k_n gelegenen Pole zu erstrecken ist, gilt, muß gezeigt werden, daß alle Pole von K einfach sind.

Wegen der Symmetrie von G in (12) läßt sich für K leicht eine Integralgleichung mit symmetrischem Kern:

$$K(xyz,\xi\eta\zeta,\lambda)\sqrt{\varrho(xyz)+j\frac{w(xyz)}{\sqrt{\lambda}}} = G(xyz,\xi\eta\zeta)\sqrt{\varrho(xyz)+j\frac{w(xyz)}{\sqrt{\lambda}}}$$
$$+\lambda\iiint G(xyz,\mathfrak{x}\mathfrak{y}\mathfrak{z})\sqrt{\varrho(xyz)+j\frac{w(xyz)}{\sqrt{\lambda}}}$$
$$\times\sqrt{\varrho(\mathfrak{x}\mathfrak{y}\mathfrak{z})+j\frac{w(\mathfrak{x}\mathfrak{y}\mathfrak{z})}{\sqrt{\lambda}}}\left\{K(\mathfrak{x}\mathfrak{y}\mathfrak{z},\xi\eta\zeta,\lambda)\cdot\sqrt{\varrho(\mathfrak{x}\mathfrak{y}\mathfrak{z})+j\frac{w(\mathfrak{x}\mathfrak{y}\mathfrak{z})}{\sqrt{\lambda}}}\right\}d\mathfrak{x}\,d\mathfrak{y}\,d\mathfrak{z}$$

anschreiben. Nach einem bekannten Satze kann jetzt

$$K(xyz,\xi\eta\zeta,\lambda)\sqrt{\varrho(xyz)+j\frac{w(xyz)}{\sqrt{\lambda}}}$$

nur einfache Pole haben. Gleiches gilt also von $K(xyz,\xi\eta\zeta,\lambda)$.

Weiter muß, um die Partialbruchzerlegung durchzuführen, bewiesen werden, daß die linke Seite der Cauchyschen Gleichung nach Null strebt für $n\to\infty$.

Hierzu genügt es, zu zeigen, daß die Entwicklung

$$K(xyz,\xi\eta\zeta,\lambda)=\sum_{i=1}^{\infty}\frac{R(\lambda_i)}{\lambda_i-\lambda}$$

für $|\lambda|\to\infty$ zu Recht besteht. Denn, da in der Cauchyschen Gleichung $|\lambda|$ innerhalb k_n liegt, ist $|\lambda|\to\infty$ identisch mit $n\to\infty$. Das *asymptotische* Bestehen der letzten Gleichung zieht also das Verschwinden der linken Seite in Cauchys Formel für $n\to\infty$ nach sich. Wir haben aber oben gezeigt, daß die Entwicklung von K für $|\lambda|\to\infty$ in (13) übergehen muß, womit die *allgemeine* Existenz der zuletzt angeschriebenen Entwicklung für K bewiesen ist[7]).

[7]) Das Verschwinden der linken Seite der Cauchyschen Gleichung für $n\to\infty$ und somit $|\mu|\to\infty$, das den Kernpunkt unserer Überlegung bildet, läßt sich einfach direkt folgendermaßen beweisen.

Für $|\mu|\to\infty$ schließt man aus Gleichung (12):

$$\lim_{|\mu|\to\infty}K(xyz\xi\eta\zeta\mu)=K^*(xyz\xi\eta\zeta\mu)\{1+O|\mu^{-\frac{1}{2}}|\},$$

wo K^* die im Text angegebene Bedeutung hat.

(Fortsetzung der Fußnote [7]) auf nächster Seite.)

III. Bilinearformeln und Entwicklungssatz.

Zur Berechnung von $R(\lambda_i)$ überlegen wir, daß nach der Fredholmschen Theorie $K(x\,y\,z,\xi\,\eta\,\zeta,\lambda)$ gegeben ist durch den Quotienten zweier ganzer transzendenter Funktionen von λ, wobei der Nenner nur von λ abhängt. Für $\lambda \to \lambda_i$ strebt dieser Nenner wie $\lambda_i - \lambda$ nach Null, woraus hervorgeht, daß das Residuum $R(\lambda_i)$ von K für $\lambda \to \lambda_i$ bestimmt wird durch die zur Gleichung (12) gehörige homogene Integralgleichung.

Da $G(x\,y\,z,\xi\,\eta\,\zeta)$ ein reeller symmetrischer Kern ist, müssen auch $K(x\,y\,z,\xi\,\eta\,\zeta,\lambda)$ und $R(x\,y\,z,\xi\,\eta\,\zeta,\lambda_i)$ symmetrisch in $x\,y\,z$ und $\xi\,\eta\,\zeta$ sein. Dann folgt aber, daß die Gleichung

$$R(x\,y\,z,\xi\,\eta\,\zeta,\lambda_i) = \lambda_i \iiint G(x\,y\,z,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,\varrho(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,R(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z},\xi\,\eta\,\zeta,\lambda_i)$$
$$\times\, d\mathfrak{x}\,d\mathfrak{y}\,d\mathfrak{z} + j\sqrt{\lambda_i}\iiint G(x\,y\,z,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,w(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z})\,R(\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z},\xi\,\eta\,\zeta,\lambda_i)\cdot d\mathfrak{x}\,d\mathfrak{y}\,d\mathfrak{z}$$

nur die Lösungen

$$R(x\,y\,z,\xi\,\eta\,\zeta,\lambda_i) = c_i\,v_i(x\,y\,z)\,v_i(\xi\,\eta\,\zeta),$$

wo v_i Eigenfunktionen von (3) sind, haben kann. Unsere noch übrigbleibende Aufgabe ist jetzt, die Koeffizienten c_i, die für jedes endliche λ_i notwendig beschränkt sind, zu berechnen.

Indessen werden wir zunächst den Entwicklungssatz erledigen und erst ganz zuletzt die Berechnung von c_i, bei der man vorteilhaft vom Entwicklungssatz Gebrauch machen kann, vornehmen.

Das Ergebnis der bisherigen Betrachtungen kann dahin zusammengefaßt werden, *daß die verallgemeinerte Greensche Funktion* $K(x\,y\,z,\xi\,\eta\,\zeta,\lambda)$ *der Gleichung* (1a) *in eine konvergente bilineare Reihe*

$$K(x\,y\,z,\xi\,\eta\,\zeta,\lambda) = \sum_i \frac{c_i\,v_i(x\,y\,z)\,v_i(\xi\,\eta\,\zeta)}{\lambda_i - \lambda} \tag{14}$$

Für K^* gilt aber wegen des Bestehens der konvergenten Bilinearreihe (13) sicher:

$$\lim_{n\to\infty} \oint_{k_n} \frac{K^*(x\,y\,z\,\xi\,\eta\,\zeta\,\mu)}{\lambda - \mu}\,d\mu = 0.$$

Folglich ist auch

$$\lim_{n\to\infty} \oint_{k_n} \frac{K(x\,y\,z\,\xi\,\eta\,\zeta\,\mu)}{\lambda - \mu}\,d\mu = \lim_{n\to\infty} \oint_{k_n} \frac{K^*(1 + O\,|\mu^{-\frac{1}{2}}|)}{\lambda - \mu}\,d\mu = 0, \quad \text{q. e. d.}$$

Ich möchte hier bemerken, daß dieses asymptotische Verschwinden des Konturintegrales den Schlüssel enthält zur Erweiterung des sogenannten Heavisideschen Theorems auf kontinuierliche Systeme, welche von T. J. I'A. Bromwich in seiner bekannten Arbeit „Normal Coordinates in Dynamical Systems“ (Proc. London Mathem. Soc. (2) **15** (1916), S. 420) auf anderem Wege versucht worden ist, aber nach seiner eigenen Angabe sogar für eine Dimension nicht streng durchgeführt werden konnte.

nach den Eigenfunktionen von (3) *entwickelt werden kann. Insbesondere besagt* (14), *daß für die Greensche Funktion* $G(x\,y\,z, \xi\,\eta\,\zeta)$ *im gewöhnlichen Sinne auch bei Eigenwertproblemen mit Dämpfung die Formel*

$$G(x\,y\,z, \xi\,\eta\,\zeta) = \sum_i \frac{c_i\, v_i(x\,y\,z)\, v_i(\xi\,\eta\,\zeta)}{\lambda_i} \tag{15}$$

gültig ist, wo v_i *und* λ_i *Eigenfunktionen und Eigenwerte von* (3) *darstellen*[8]).

Wenn man für die Gleichung (1a) mit $w = 0$ beweisen könnte, daß (13) und die hieraus für $\lambda = 0$ hervorgehende Bilinearreihe für G *gleichmäßig* konvergieren, würde durch unsere Überlegungen auch die gleichmäßige Konvergenz von (14) und (15) folgen. Da aber die gleichmäßige Konvergenz von (13) für mehr als eine Dimension *nicht* bewiesen ist, können wir auch nur auf die gleichmäßige Konvergenz von (14) und (15) für eine Dimension schließen[9]).

Die Aufgabe, eine Funktion F nach den Eigenfunktionen von (3) zu entwickeln, läßt sich lösen, wenn für F die Operation:

$$L[F] + \lambda\,\varrho\,F + j\sqrt{\lambda}\,w\,F = -f$$

zu einer Funktion f führt, die, mit einer Eigenfunktion v_i von (3) multipliziert und über das ganze Gebiet integriert, für jedes i ein beschränktes Ergebnis liefert:

$$\left| \iiint f(x\,y\,z)\, v_i(x\,y\,z)\, dx\, dy\, dz \right| < M.$$

Dies ist insbesondere erfüllt, wenn f selber beschränkt, also F zweimal stetig nach $x\,y\,z$ differenzierbar ist.

Für eine solche Funktion F kann man mit Hilfe von K schreiben:

$$F(x\,y\,z) = \iiint K(x\,y\,z, \xi\,\eta\,\zeta, \lambda)\, f(\xi\,\eta\,\zeta)\, d\xi\, d\eta\, d\zeta = \sum_i \frac{c_i\, f_i\, v_i(x\,y\,z)}{\lambda_i - \lambda}, \tag{17}$$

mit

$$f_i = \iiint f(x\,y\,z)\, v_i(x\,y\,z)\, dx\, dy\, dz.$$

Die Rechtfertigung der gliedweisen Integration von (14) muß, da die gleichmäßige Konvergenz von (14) nicht feststeht, indirekt stattfinden, dadurch, daß die gleichmäßige Konvergenz von (17) bewiesen wird.

---

[8]) Für mehr als eine Dimension konvergieren die Reihen (14) und (15) nicht mehr für $x = \xi$, $y = \eta$, $z = \zeta$.

[9]) Wohl beweist man leicht, daß die Lösung von (12) sich in eine absolut und gleichmäßig konvergente Reihe:

$$K(x\,y\,z, \xi\,\eta\,\zeta, \lambda) = G(x\,y\,z, \xi\,\eta\,\zeta) + \lambda \sum_i \frac{c_i\, v_i(x\,y\,z)\, v_i(\xi\,\eta\,\zeta)}{\lambda_i(\lambda_i - \lambda)}$$

entwickeln läßt, wo v_i und λ_i Eigenfunktionen und Eigenwerte von (3) sind.

Hierzu überlegen wir wieder, daß die Eigenfunktionen v_i und Eigenwerte λ_i von (1a) für $i \to \infty$ in jene von (4) übergehen mit einem Fehler der Ordnung $|\lambda^{-\frac{1}{2}}|$. Die Reihe:

$$F = \sum_i \frac{f_i^* v_i^*(xyz)}{\lambda_i^* \iiint \varrho (v_i^*)^2 \, dx\, dy\, dz}, \tag{17a}$$

wo

$$L[F] = -f^*; \qquad F = \iiint G(xyz, \xi\eta\zeta) f^*(\xi\eta\zeta)\, d\xi\, d\eta\, d\zeta;$$

$$f_i^* = \iiint f^*(xyz)\, v_i^*(xyz)\, dx\, dy\, dz$$

und v_i^* und λ_i^* Eigenfunktionen und Eigenwerte von (4) sind, konvergiert absolut und gleichmäßig. Hieraus schließen wir auf die absolute und gleichmäßige Konvergenz von (17) für jedes endliche λ. Andrerseits ist (17) nichts anderes als die Cauchysche Partialbruchzerlegung von F, so daß wir bewiesen haben, *daß jede zweimal stetig differenzierbare Funktion F in eine absolut und gleichmäßig konvergente Reihe nach den Eigenfunktionen v_i von* (1a), (2) *entwickelt werden kann gemäß* (17).

In Fabers früher erwähnter Arbeit steht in (17) in unsrer Bezeichnung $\sqrt{\lambda_i} - \sqrt{\lambda}$ im Nenner und dementsprechend etwas andere c_i, die aber λ nicht enthalten. Obwohl Fabers Form der unsrigen äquivalent ist, erlaubt sie nicht, den Sabineschen Satz unmittelbar einzusehen, im Gegensatz zur vorliegenden.

Die Gleichung (17) genügt, die in der mathematischen Physik bei Anfangsaufgaben und linearen nicht homogenen Randaufgaben (erzwungene Schwingung) auftretenden Entwicklungsfragen zu erledigen,

Bei der Berechnung von c_i, die wir zum Schluß dieses Abschnittes vornehmen, machen wir von (16) Gebrauch und schreiben v_i an die Stelle von F:

$$L[v_i] + \lambda \varrho v_i + j\sqrt{\lambda}\, w v_i = -f = (\lambda - \lambda_i)\varrho v_i + j(\sqrt{\lambda} - \sqrt{\lambda_i})\, w v_i,$$

wodurch in (17) die Größen f_i

$$f_i = (\lambda_i - \lambda)\left\{\iiint \varrho v_i^2 \, dx\, dy\, dz + \frac{j}{\sqrt{\lambda} + \sqrt{\lambda_i}} \iiint w v_i^2 \, dx\, dy\, dz\right\}$$

werden. Hiernach lassen wir λ sich dem Werte λ_i unbeschränkt nähern:

$$\lambda \to \lambda_i,$$

wodurch alle f_i nach Null streben. Aus (17) erhellt, daß alle Reihenglieder verschwinden, außer dem Gliede mit dem Zeiger i, so daß (17) sich schreibt:

$$v_i = c_i v_i \left\{\iiint \varrho v_i^2 \, dx\, dy\, dz + \frac{j}{2\sqrt{\lambda_i}} \iiint w v_i^2 \, dx\, dy\, dz\right\}. \tag{18}$$

Aus (18) geht sofort hervor:

$$(19) \qquad c_i = \left\{ \iiint \varrho v_i^2 \, dx\, dy\, dz + \frac{j}{2\sqrt{\lambda_i}} \iiint w v_i^2 \, dx\, dy\, dz \right\}^{-1},$$

womit der noch fehlende Koeffizient in der Entwicklung (14) gefunden ist.

Die entstehende Reihe (14), (19) erlaubt, nachträglich den bereits früher aus (12) gezogenen Schluß zu verifizieren, nachdem die Entwicklung (14) von K asymptotisch in die Entwicklung (13) von K^* übergeht.

IV. Über die Akustik großer Räume.

Experimentell werden die akustischen Verhältnisse von Räumen oft durch die Nachhalldauer charakterisiert. Diese ist gegeben durch die Zeit, welche die Schallintensität in einem bis zum stationären Zustand mit einem Ton, dessen Wellenlänge klein ist gegenüber den Raumdimensionen, angeblasenen Raume braucht, um bis zur Hörbarkeitsschwelle herabzusinken. Die Messungen werden dahin reduziert, daß im Ergebnis die Anfangsintensität ein bestimmtes Vielfaches, etwa das Millionenfache der Schwellenintensität beträgt. Sabines in der Einleitung erwähnte Satz besagt, daß die so gemessene Nachhalldauer in genügend großen Räumen nur von der gesamten Oberfläche des dämpfenden Materiales, jeweils mit einer Dämpfungskonstanten multipliziert, und dem Volumen abhängt, und zwar der zuerst genannten Größe umgekehrt, der zuletzt genannten direkt proportional ist. Diesen Satz beabsichtige ich hier im Limes, für ein Verhältnis der erregenden Wellenlänge zu Raumdimensionen, das nach Null strebt, abzuleiten.

In Räumen ist bei akustischen Problemen, wie eben erwähnt, die Dämpfung nicht stetig über das ganze Gebiet verteilt, sondern nur an der Gebietsbegrenzung vorhanden. Unsre erste Aufgabe ist, zu untersuchen, wie eine solche Dämpfungsverteilung sich in unsren Formeln kundtut.

Zunächst zeigt die Betrachtung etwa der Courantschen Ableitung von (6), (8) und (10), daß ein Parameter in den Randbedingungen, also ein Abändern von (2) in irgendwelche andere homogene Randbedingungen, keinen Einfluß auf diese asymptotische Schätzung hat.

Folglich kann auch die Ableitung von (7), (9) und (11) in derselben Weise wie oben durchgeführt werden und muß nachträglich in diesen Formeln durch einen Grenzprozeß, der Poincarés „Méthode de Balayage" für Potentialaufgaben nachgebildet ist, dem Umstand Rechnung getragen werden, daß w im Raume überall verschwindet, während doch β endlich bleibt.

Man kann dies einfach dadurch erreichen, daß etwa das Raumintegral im Zähler von (11) auf ein zwischen der Grenze und einer sich der Grenze

eng anschmiegenden Fläche eingeschlossenes Gebiet erstreckt wird. Ist dn das Längenelement normal zu diesen Flächen, df das Element der Flächen selber, so schreibt sich das genannte Integral:

$$\iint df \int dn \frac{w}{p} \sqrt{\frac{\varrho}{p}}.$$

Läßt man die zwei Flächen genügend eng zusammenrücken, so sind im Intervall senkrecht zu und zwischen diesen Flächen ϱ und p konstant. Wir machen nun auch w in diesem Intervall konstant und gleich

$$\frac{r}{\varepsilon},$$

wenn ε den senkrechten Abstand bei den Flächen bezeichnet[10]. Durch diese Schreibweise findet man an Stelle von (11):

$$\text{(11a)} \qquad \lim_{\alpha \to \infty} \beta = -\frac{1}{2} \frac{\iint \frac{r}{p} \sqrt{\frac{\varrho}{p}}\, df}{\iiint \left(\frac{\varrho}{p}\right)^{2} dx\, dy\, dz} + O\left(\frac{\ln \alpha}{\alpha}\right).$$

Analoge Gleichungen wie (11a) lassen sich im ein- und zweidimensionalen Fall anschreiben.

Schließlich möchte ich noch bemerken, daß w und r beliebige *beschränkte* für großes $|\lambda|$ schwach veränderliche Funktionen von λ sein können, ohne daß sich an den vorhergegangenen Überlegungen etwas Wesentliches ändert. Hierdurch sind wir im akustischen Problem in der Lage, irgendeine passende Annahme über die Art der Dämpfung zu machen; z. B. fällt die Dämpfung an porösen Wänden in den Bereich unsrer Formeln.

Zum Beweis des Sabineschen Satzes ist es zunächst notwendig, die stationäre in einem Raume sich unter der Einwirkung einer rein periodischen Schallquelle einstellende Amplitude des Geschwindigkeitspotentiales Φ zu berechnen. Hierzu haben wir die Differentialgleichung:

$$L[\Phi] + \lambda \varrho \Phi + j \sqrt{\lambda}\, w \Phi = -F,$$

wo F die Quellenverteilung repräsentiert und w das oben besprochene Verhalten zeigt, etwa unter der Randbedingung (2) zu lösen. Die stationäre Φ-Amplitude ergibt sich hierbei aus der Entwicklungsformel (17) des vorigen Abschnittes zu

$$\Phi = \sum_i \frac{c_i v_i(xyz)}{\lambda_i - \lambda} \iiint F v_i\, dx\, dy\, dz.$$

[10] Dieses Unendlichwerden von w hat auf die Schätzung (3a) keinen Einfluß, da dort nur das *Integral* über w in Frage kommt.

Zieht man noch die Zeitabhängigkeit des Geschwindigkeitspotentiales sowie der erregenden Quellenverteilung:

$$\Phi_0 = \Phi e^{-j\sqrt{\lambda}t},$$
$$F_0 = F e^{-j\sqrt{\lambda}t}$$

in Betracht, so ergibt sich für Φ_0 und $\frac{\partial \Phi_0}{\partial t}$ zur Zeit $t = t_0$:

$$(20)\quad \left\{\begin{aligned} (\Phi_0)_{t=t_0} &= e^{-j\sqrt{\lambda}t_0}\sum_i \frac{c_i \iiint F v_i\, dx\, dy\, dz}{\lambda_i - \lambda} v_i(xyz);\\ \left(\frac{\partial \Phi_0}{\partial t}\right)_{t=t_0} &= -j\sqrt{\lambda}\, e^{-j\sqrt{\lambda}t_0}\sum_i \frac{c_i F_i}{\lambda_i - \lambda} v_i(xyz).\end{aligned}\right.$$

Zur Zeit $t = t_0$ soll die erregende Quelle ihre Tätigkeit einstellen. Dann bestimmt sich das Abklingen von Φ_0 in der Folgezeit aus

$$(21)\quad \Phi_0 = \sum_i A_i v_i(xyz)\, e^{(\beta_i + j\alpha_i)(t-t_0)} + \sum_i B_i v_i^*(xyz)\, e^{(\beta_i - j\alpha_i)(t-t_0)}.$$

(Stern bedeutet konjugiert komplexen Wert.)

Die Koeffizienten A_i und B_i sind aus den Anfangsbedingungen (20) zu berechnen.

Mein Ziel ist, zu zeigen, daß alle A_i und B_i unterhalb einem gewissen, sehr großen Indexwert i beliebig wenig von Null verschieden sind, falls $|\lambda| \to \infty$ wächst. Wenn dies der Fall ist, zeigt (21), daß das Abklingen von Φ_0, da β_i für $i \to \infty$ von i unabhängig wird und durch (11) bzw. (11a) gegeben ist, in einem durch die zuletzt genannten Gleichungen bestimmten Tempo stattfindet, d. h. nur vom Raumvolumen und vom Dämpfungsintegral in der vom Sabineschen Satze geforderten Weise abhängt. Andrerseits ist klar, daß A_i und B_i die eben genannte Bedingung erfüllen, wenn in beiden Gleichungen (20) die Koeffizienten von v_i unterhalb einem gewissen, sehr großen i absolut beliebig klein werden für $|\lambda| \to \infty$. Zunächst ist hierzu notwendig, von der Quellenverteilung F zu fordern, daß die Integrale

$$\left|\iiint F(xyz)\, v_i(xyz)\, dx\, dy\, dz\right|$$

für beliebiges i unterhalb einer festen Schranke M bleiben. Unter dieser Voraussetzung erfüllen aber, wie sofort ersichtlich, die Koeffizienten von v_i in (20) die eben genannte Bedingung, denn

$$\left|\frac{1}{\lambda_i - \lambda}\right| \quad \text{bzw.} \quad \left|\frac{\sqrt{\lambda}}{\lambda_i - \lambda}\right|$$

weichen für $|\lambda| \to \infty$ bis zu einem festen, sehr großen i beliebig wenig von Null ab.

Der Sabinesche Satz ist hiermit asymptotisch als richtig nachgewiesen.

Es ist klar, daß sich dem Sabineschen nachgebildete Sätze in der Elektrodynamik und der Mechanik formulieren lassen. Allgemein kann man sagen:

Wird ein in bezug auf „elastische“ Eigenschaften homogenes Kontinuum, dessen freie Schwingungen durch (1) *beschrieben werden, bis zum stationären Zustand durch eine periodische Quelle zu Schwingungen angeregt, derart, daß die größte Wellenlänge noch klein zur größten Eigenwellenlänge des Kontinuums ist, so ist die vom Augenblick des Aussetzens der Quelle gemessene Abklingungszeit der gesamten Dämpfung umgekehrt, dem Volumen direkt proportional und unabhängig von der Gestalt des Kontinuums, der Art der Quelle, des Ortes der Messung und von dem erregenden Tone.*

Der Satz läßt sich auch beweisen für Kontinua mit Gleichungen vierter Ordnung (Platte).

Ein ähnlicher Satz gilt beim Wecken eines Systems aus dem Ruhezustande durch eine periodische Quelle hoher Frequenz. Der Anlaufvorgang ist dabei dem Abklingungsvorgang komplementär.

Eindhoven (Niederlande), Januar 1929,
Natuurkundig Laboratorium der N. V. Philips' Gloeilampenfabrieken.

(Eingegangen am 22. 2. 1929.)

Auf die Riemann-Metrik und den Fern-Parallelismus gegründete einheitliche Feldtheorie.

Von

A. Einstein in Berlin.

In der folgenden Arbeit soll die von mir seit einem Jahre entwickelte Theorie so dargestellt werden, daß sie von jedem Kenner der allgemeinen Relativitätstheorie bequem verstanden werden kann. Die folgende Darstellung ist nötig, weil die Lektüre der früheren Arbeiten durch seither gefundene Zusammenhänge und Verbesserungen unnötigen Zeitverlust bedingen würde. Der Gegenstand ist so dargestellt, wie es mir für ein bequemes Eindringen am zweckmäßigsten erscheint. Insbesondere durch die Herren Weitzenböck und Cartan erfuhr ich, daß die Behandlung von Kontinua der hier in Betracht kommenden Gattung an sich nicht neu sei. Herr Cartan war so freundlich, eine Abhandlung über die Geschichte des hier in Betracht kommenden mathematischen Gegenstandes zu verfassen, damit durch sie der Inhalt meiner Darstellung ergänzt werde; dieselbe ist in dieser Zeitschrift unmittelbar nach der meinigen abgedruckt. Ich sage Herrn Cartan auch an dieser Stelle für seinen wertvollen Beitrag meinen besten Dank. Was an der vorliegenden Abhandlung das Wichtigste und jedenfalls neu ist, das ist die Auffindung der einfachsten Feldgesetze, welchen eine Riemannsche Mannigfaltigkeit mit Fern-Parallelismus unterworfen werden kann. Auf die Frage der physikalischen Bedeutung der Theorie gehe ich nur kurz ein.

§ 1.

Die Struktur des Kontinuums.

Da die Dimensionszahl in den folgenden Überlegungen keine Rolle spielt, legen wir ein n-dimensionales Kontinuum zugrunde. Um den Tatsachen der Metrik und der Gravitation gerecht zu werden, nehmen wir die Existenz einer Riemann-Metrik an. In der Natur existieren aber auch

die elektromagnetischen Felder, welche durch die Riemann-Metrik nicht dargestellt werden können. Es entsteht die Frage: Wie können wir unserem Riemannschen Raume in logisch natürlicher Weise noch eine weitere Struktur zuschreiben, derart, daß das Ganze einheitlichen Charakter hat?

Das Kontinuum ist in der Umgebung jedes Punktes P (pseudo-)euklidisch. Es existiert in jedem Punkte ein lokales kartesisches Koordinatensystem (bzw. orthogonales n-Bein), in bezug auf welches der pythagoreische Satz gilt. Die Orientierung dieser n-Beine spielt in einer Riemannschen Mannigfaltigkeit keine Rolle. Wir wollen nun annehmen, daß zwischen diesen elementaren euklidischen Räumen auch noch eine Richtungsbeziehung herrscht. Wir wollen annehmen, daß es vermöge der Raumstruktur wie in der euklidischen Geometrie sinnvoll sei, von einer Parallel-Orientierung sämtlicher lokaler n-Beine zu sprechen (was in einem Raume mit *nur* metrischer Struktur sinnlos wäre). Wir denken uns die orthogonalen n-Beine im folgenden stets in Parallel-Orientierung. Die an sich willkürliche Orientierung des lokalen n-Beines in *einem* Punkte P bestimmt dann die Orientierung der lokalen n-Beine in allen Punkten des Kontinuums eindeutig. Unsere Aufgabe besteht nun zunächst darin, ein solches Kontinuum mathematisch zu beschreiben, sodann aber darin, die einfachsten beschränkenden Gesetze aufzustellen, welchen ein solches Kontinuum unterworfen werden kann. Wir tun dies in der Hoffnung, so die allgemeinen Naturgesetze abzuleiten, wie die frühere allgemeine Relativitätstheorie dies für die Gravitation unter Zugrundelegung einer bloß metrischen Raumstruktur versucht hat.

§ 2.

Mathematische Beschreibung der Raumstruktur.

Das lokale n-Bein besteht aus n aufeinander senkrechten Einheitsvektoren, deren auf ein beliebiges Gaußsches Koordinatensystem bezogene Komponenten h_s^{ν} seien. Hier wie immer soll ein unterer lateinischer Index die Zugehörigkeit zu einem bestimmten Bein des n-Beins, ein griechischer Index je nach seiner oberen oder unteren Stellung den kontravarianten bzw. kovarianten Transformationscharakter der betreffenden Größe bezüglich einer Änderung des Gaußschen Koordinatensystems zum Ausdruck bringen.

Die allgemeine Transformationseigenschaft der h_s^{ν} ist folgende. Dreht man alle Lokalsysteme bzw. n-Beine in gleicher Weise, was ja erlaubt ist, und führt man zugleich ein neues Gaußsches Koordinatensystem ein, so besteht zwischen den neuen und alten h_s^{ν} das Transformationsgesetz

$$h_s^{\nu'} = \alpha_{st} \frac{\partial x^{\nu'}}{\partial x^{\alpha}} h_t^{\alpha}, \tag{1}$$

wobei die konstanten Koeffizienten α_{st} ein orthogonales System bilden:

$$(2) \qquad \alpha_{sa}\alpha_{sb} = \alpha_{as}\alpha_{bs} = \delta_{ab} = \begin{cases} 1, & \text{wenn } a = b, \\ 0, & \text{wenn } a \neq b. \end{cases}$$

Das Transformationsgesetz (1) läßt sich ohne weiteres auf Gebilde verallgemeinern, deren Komponenten beliebig viele lokale und Koordinaten-Indizes haben. Solche Gebilde nennen wir Tensoren. Hieraus ergeben sich die algebraischen Gesetze der Tensoren unmittelbar (Addition, Multiplikation, Verjüngung nach lateinischen und griechischen Indizes).

Die $h_s{}^\nu$ nennen wir die Komponenten des Fundamentaltensors. Hat ein Vektor im Lokalsystem die Komponenten A_s, bezüglich des Gaußschen Systems die Koordinaten A^ν, so gilt nach der Bedeutung der $h_s{}^\nu$:

$$(3) \qquad A^\nu = h_s{}^\nu A_s$$

oder — nach den A_s aufgelöst —

$$(4) \qquad A_s = h_{s\nu} A^\nu.$$

Der Tensorcharakter der normierten Unterdeterminanten $h_{s\nu}$ der $h_s{}^\nu$ ergibt sich aus (4). $h_{s\nu}$ sind die kovarianten Komponenten des Fundamentaltensors. Zwischen den $h_{s\nu}$ und $h_s{}^\nu$ gelten die Beziehungen

$$(5) \qquad h_{s\mu} h_s{}^\nu = \delta_\mu{}^\nu \begin{cases} = 1, & \text{wenn } \mu = \nu, \\ = 0, & \text{wenn } \mu \neq \nu, \end{cases}$$

$$(6) \qquad h_{s\mu} h_t{}^\mu = \delta_{st}.$$

Wegen der Orthogonalität des Lokalsystems gilt für den Betrag des Vektors

$$(6) \qquad A^2 = A_s^2 = h_{s\mu} h_{s\nu} A^\mu A^\nu = g_{\mu\nu} A^\mu A^\nu;$$

also sind

$$(7) \qquad g_{\mu\nu} = h_{s\mu} h_{s\nu}$$

die Koeffizienten der Metrik.

Der Fundamentaltensor gestattet (vgl. (3) und (4)), lokale Indizes in Koordinatenindizes zu verwandeln und umgekehrt (durch Multiplikation und Verjüngung), so daß es nur eine Formfrage bedeutet, mit Tensoren welchen Indexcharakters man operieren will.

Es ist klar, daß auch die Beziehungen gelten

$$(3\text{a}) \qquad A_\nu = h_{s\nu} A_s,$$

$$(4\text{a}) \qquad A_s = h_s{}^\nu A_\nu.$$

Es gilt ferner die Determinanten-Relation

$$(8) \qquad g = |g_{\sigma\tau}| = |h_{a\sigma}|^2 = h^2,$$

so daß die Invariante des Volumelementes $\sqrt{g}\, d\tau$ die Form $h\, d\tau$ annimmt.

Dem besonderen Charakter der Zeit wird bei unserem 4-dimensionalen Kontinuum von Raum und Zeit am bequemsten dadurch Rechnung getragen, daß die x^4-Koordinate (sowohl lokal als allgemein) rein imaginär genommen wird und ebenso alle Tensorkomponenten mit einer ungeraden Zahl von Indizes 4.

§ 3.

Differentialrelationen.

Bezeichnen wir mit δ den Zuwachs, den die Komponenten eines Vektors oder Tensors bei „Parallelverschiebung" im Sinne von Levi-Civita beim Übergang zu einem unendlich benachbarten Punkte des Kontinuums erfahren, so ist nach dem obigen

$$(9) \qquad 0 = \delta A_s = \delta (h_{s\alpha} A^\alpha) = \delta (h_s{}^\alpha A_\alpha).$$

Die Auflösung der Klammern ergibt

$$h_{s\alpha} \delta A^\alpha + A^\alpha h_{s\alpha,\beta} \delta x^\beta = 0,$$
$$h_s{}^\alpha \delta A_\alpha + A_\alpha h_s{}^\alpha{}_{,\beta} \delta x^\beta = 0,$$

wobei das Komma im zweiten Gliede gewöhnliche Differentiation nach x^β bedeutet. Durch Auflösen dieser Gleichungen erhält man

$$(10) \qquad \delta A^\sigma = - A^\alpha \Delta_\alpha{}^\sigma{}_\beta \delta x^\beta,$$

$$(11) \qquad \delta A_\sigma = A_\alpha \Delta_\sigma{}^\alpha{}_\beta \delta x^\beta,$$

wobei gesetzt ist

$$(12) \qquad \Delta_\alpha{}^\sigma{}_\beta = h_s{}^\sigma h_{s\alpha,\beta} = - h_{s\alpha} h_s{}^\sigma{}_{,\beta}.$$

(Die letzte Umformung beruht auf (5).)

Dies Gesetz der Parallelverschiebung ist im Gegensatz zur Riemann-Geometrie im allgemeinen nicht symmetrisch. Ist es symmetrisch, so besteht die euklidische Geometrie; denn man hat dann

$$\Delta_\alpha{}^\sigma{}_\beta - \Delta_\beta{}^\sigma{}_\alpha = 0$$

oder

$$h_{s\alpha,\beta} - h_{s\beta,\alpha} = 0.$$

Dann ist aber

$$h_{s\alpha} = \frac{\partial \psi_s}{\partial x_\alpha}.$$

Wählt man die ψ_s als neue Variable x_s', so ist

$$(13) \qquad h_{s\alpha} = \delta_{s\alpha},$$

was die Behauptung beweist.

Kovariante Differentiation. Die lokalen Komponenten A_s eines Vektors sind invariant bezüglich einer beliebigen Koordinatentransformation. Daraus folgt sofort der Tensorcharakter der Differentialquotienten

(14) $$A_{s,\alpha}.$$

Ersetzt man dies auf Grund von (4a) durch

$$(h_s{}^\sigma A_\sigma)_{,\alpha},$$

so ergibt sich hieraus der Tensorcharakter von

$$h_s{}^\sigma A_{\sigma,\alpha} + A_\sigma h_s{}^\sigma{}_{,\alpha},$$

also auch (nach Multiplikation mit $h_{s\tau}$) von

$$A_{\tau,\alpha} + A_\sigma h_s{}^\sigma{}_{,\alpha} h_{s\tau}$$

und von

$$A_{\tau,\alpha} - A_\sigma h_s{}^\sigma h_{s\tau,\alpha}$$

oder gemäß (16) von

$$A_{\tau,\alpha} - A_\sigma \Delta_\tau{}^\sigma{}_\alpha.$$

Dies bezeichnen wir als kovariante Ableitung ($A_{\tau;\alpha}$) von A_τ.

Wir haben so als Gesetz der kovarianten Differentiation

(15) $$A_{\sigma;\tau} = A_{\sigma,\tau} - A_\alpha \Delta_\sigma{}^\alpha{}_\tau$$

gewonnen. Analog folgt aus (3) auch die Formel

(16) $$A^\sigma{}_{;\tau} = A^\sigma{}_{,\tau} + A^\alpha \Delta_\alpha{}^\sigma{}_\tau.$$

Nun ergibt sich analog das kovariante Differentiationsgesetz für beliebige Tensoren. Wir erläutern es durch das Beispiel:

(17) $$A_\alpha{}^\sigma{}_{\tau;\varrho} = A_\alpha{}^\sigma{}_{\tau,\varrho} + A_\alpha{}^\alpha{}_\tau \Delta_\alpha{}^\sigma{}_\varrho - A_\alpha{}^\sigma{}_\alpha \Delta_\tau{}^\alpha{}_\varrho.$$

Da sich vermittels des Fundamentaltensors $h_s{}^\alpha$ lokale (lateinische) Indizes in Koordinaten- (griechische) Indizes umwandeln lassen, so steht es frei, bei der Formulierung irgendwelcher Tensorbeziehungen die lokalen oder die Koordinaten-Indizes zu bevorzugen. Ersterer Weg wurde von den italienischen Kollegen (Levi-Civita, Palatini) bevorzugt, während ich mich vorzugsweise der Koordinatenindizes bedient habe.

Divergenz. Durch Verjüngung des kovarianten Differentialquotienten erhält man wie in dem auf die Metrik allein gegründeten absoluten Differentialkalkül die Divergenz. Z. B. erhält man aus (21) durch Verjüngung nach den Indizes σ und ϱ den Tensor

$$A_{\alpha\tau} = A_\alpha{}^\sigma{}_{\tau;\sigma}.$$

In früheren Arbeiten habe ich noch andere Divergenzoperatoren eingeführt, bin aber davon abgekommen, jenen Operatoren eine besondere Bedeutung zuzuschreiben.

Kovariante Differentialquotienten des Fundamentaltensors. Man findet aus den abgeleiteten Formeln leicht, daß die kovarianten Ableitungen und Divergenzen des Fundamentaltensors verschwinden. Es ist z. B.

$$
\begin{aligned}
(18) \qquad h_{s\ ;\tau}^{\ \nu} &\equiv h_{s\ ,\tau}^{\ \nu} + h_s^{\ \alpha} \Delta_{\alpha\ \tau}^{\ \nu} \equiv \delta_{st}(h_{t\ ,\tau}^{\ \nu} + h_t^{\ \alpha} \Delta_{\alpha\ ,\tau}^{\ \nu}) \\
&\equiv h_s^{\ \alpha}(h_{t\alpha} h_{t\ ,\tau}^{\ \nu} + \Delta_{\alpha\ \tau}^{\ \nu}) \equiv h_s^{\ \alpha}(-\Delta_{\alpha\ \tau}^{\ \nu} + \Delta_{\alpha\ \tau}^{\ \nu}) \equiv 0.
\end{aligned}
$$

Analog beweist man auch

$$
(18\text{a}) \qquad h_{s\ ;\tau}^{\ \nu} \equiv g^{\mu\nu}_{\ \ ;\tau} \equiv g_{\mu\nu;\tau} \equiv 0.
$$

Ebenso verschwinden natürlich die Divergenzen $h_{s\ ;\nu}^{\ \nu}$ und $g^{\mu\nu}_{\ \ ;\nu}$.

Differentiation von Tensorprodukten. Wie in dem geläufigen Differentialkalkül läßt sich der kovariante Differentialquotient eines Tensorproduktes aus den Differentialquotienten der Faktoren ausdrücken. Sind $S_{\cdot}^{\ \cdot}$ und $T_{\cdot}^{\ \cdot}$ Tensoren von beliebigem Indexcharakter, so ist

$$
(19) \qquad (S_{\cdot}^{\ \cdot} T_{\cdot}^{\ \cdot})_{;\alpha} = S_{\cdot\ ;\alpha}^{\ \cdot} T_{\cdot}^{\ \cdot} + T_{\cdot\ ;\alpha}^{\ \cdot} S_{\cdot}^{\ \cdot}.
$$

Hieraus und aus dem Verschwinden des kovarianten Differentialquotienten des Fundamentaltensors folgt, daß man diesen nach Belieben mit dem Differentiationszeichen (;) vertauschen kann.

„Krümmung". Aus der Hypothese des „Fern-Parallelismus" bzw. aus Gleichung (9) ergibt sich die Integrabilität des Verschiebungsgesetzes (10) bzw. (11). Daraus folgt

$$
(20) \qquad 0 \equiv -\Delta_{\varkappa\ \lambda;\mu}^{\ \iota} \equiv -\Delta_{\varkappa\ \lambda,\mu}^{\ \iota} + \Delta_{\varkappa\ \mu,\lambda}^{\ \iota} + \Delta_{\sigma\ \lambda}^{\ \iota} \Delta_{\varkappa\ \mu}^{\ \sigma} - \Delta_{\sigma\ \mu}^{\ \iota} \Delta_{\varkappa\ \lambda}^{\ \sigma}.
$$

Diese Bedingungen müssen die Δ erfüllen, damit sie sich gemäß (12) durch die Größen h ausdrücken lassen. Man sieht aus (20), daß die gesetzliche Charakterisierung bei einer Mannigfaltigkeit der hier betrachteten Art eine ganz andersartige sein muß als gemäß der früheren Theorie. Zwar existieren gemäß der neuen Theorie alle Tensoren der früneren Theorie, im besonderen auch der aus den Christoffel-Symbolen gebildete Riemannsche Krümmungstensor. Aber es existieren gemäß der neuen Theorie einfachere und näherliegende tensorielle Bildungen, welche bei der Formulierung der Feldgesetze herangezogen werden können.

Der Tensor Λ. Differenzieren wir einen Skalar φ zweimal kovariant, so erhalten wir gemäß (15) den Tensor

$$
\varphi_{,\sigma,\tau} - \varphi_{,\alpha} \Delta_{\sigma\ \tau}^{\ \alpha}.
$$

Durch Vertauschen von σ und τ entsteht ein neuer Tensor und durch Subtraktion beider der Tensor

$$
\frac{\partial \varphi}{\partial x_\alpha} (\Delta_{\sigma\ \tau}^{\ \alpha} - \Delta_{\tau\ \sigma}^{\ \alpha}).
$$

Hieraus folgt sofort der Tensorcharakter von

$$\Lambda_{\sigma}{}^{\alpha}{}_{\tau} = \Delta_{\sigma}{}^{\alpha}{}_{\tau} - \Delta_{\tau}{}^{\alpha}{}_{\sigma}. \tag{21}$$

Es existiert also gemäß dieser Theorie ein Tensor, der nur die Komponenten $h_{s\alpha}$ des Fundamentaltensors und dessen erste Differentialquotienten enthält. Daß sein Verschwinden die Gültigkeit der euklidischen Geometrie zur Folge hat, wurde schon früher bewiesen (vgl. (13)). Eine naturgemäße gesetzliche Bestimmung eines derartigen Kontinuums wird also in Bedingungen für diesen Tensor bestehen.

Durch Verjüngen des Tensors Λ entsteht der Vektor

$$\varphi_{\sigma} = \Lambda_{\sigma}{}^{\alpha}{}_{\alpha}, \tag{22}$$

von dem ich früher annahm, daß er in dieser Theorie die Rolle des elektromagnetischen Potentials spiele. Von dieser Auffassung bin ich aber neuerdings abgekommen.

V e r t a u s c h u n g s r e g e l d e r D i f f e r e n t i a t i o n. Differentiiert man einen beliebigen Tensor $T_{\cdot}^{\cdot}$ zweimal kovariant, so gilt die wichtige Vertauschungsregel

$$T_{\cdot}^{\cdot}{}_{;\sigma;\tau} - T_{\cdot}^{\cdot}{}_{;\tau;\sigma} \equiv - T_{\cdot}^{\cdot}{}_{;\alpha} \Lambda_{\sigma}{}^{\alpha}{}_{\tau}. \tag{23}$$

B e w e i s. Ist T ein Skalar (Tensor ohne griechischen Index), so folgt der Satz mühelos aus (15). Auf diesen Spezialfall wollen wir den allgemeinen Beweis des Satzes gründen.

Wir bemerken hierzu zunächst, daß es gemäß der hier behandelten Theorie Parallel-Vektorfelder gibt. Es sind dies Vektorfelder, welche in allen Lokalsystemen dieselben Komponenten haben. Ist (a^{α}) bzw. (a_{α}) ein solches Vektorfeld, so erfüllt es die Bedingung

$$a^{\alpha}{}_{;\sigma} = 0 \quad \text{bzw.} \quad a_{\alpha;\sigma} = 0,$$

wie leicht zu beweisen ist.

Unter Benutzung solcher Parallel-Vektorfelder läßt sich die Vertauschungsregel ohne Mühe auf diejenige für einen Skalar zurückführen. Wir führen der einfachen Schreibweise wegen den Beweis für einen Tensor T^{λ} mit nur einem Index. Ist φ ein Skalar, so folgt aus den Definitionsgleichungen (16) und (21) zunächst

$$\varphi_{;\sigma;\tau} - \varphi_{;\tau;\sigma} \equiv - \varphi_{;\alpha} \Lambda_{\sigma}{}^{\alpha}{}_{\tau}.$$

Setzen wir in diese Gleichung für φ den Skalar $a_{\lambda} T^{\lambda}$ ein, wobei a_{λ} ein Parallel-Vektorfeld ist, so läßt sich bei jeder kovarianten Differentiation a_{λ} mit dem Differentiationszeichen vertauschen, so daß a_{λ} in allen Gliedern als Faktor erscheint. Man erhält also

$$[T^{\lambda}{}_{;\sigma;\tau} - T^{\lambda}{}_{;\tau;\sigma} + T^{\lambda}{}_{;\alpha} \Lambda_{\sigma}{}^{\alpha}{}_{\tau}] a_{\lambda} \equiv 0.$$

Da diese Identität für beliebige Wahl von a_λ an einer ins Auge gefaßten Stelle bestehen muß, so folgt das Verschwinden der eckigen Klammer, womit der Beweis geliefert ist. Die Verallgemeinerung auf Tensoren mit beliebig vielen griechischen Indizes liegt auf der Hand.

Identitäten für den Tensor Λ. Addiert man die drei Identitäten, welche aus (20) durch zyklische Vertauschung von $\varkappa, \lambda, \mu$ hervorgehen, so ergibt sich durch passende Zusammenfassung der Glieder mit Rücksicht auf (21) zunächst

$$0 \equiv (\Lambda_{\varkappa}{}^{\iota}{}_{\lambda,\mu} + \Lambda_{\lambda}{}^{\iota}{}_{\mu,\varkappa} + \Lambda_{\mu}{}^{\iota}{}_{\varkappa,\lambda}) + (\Delta_{\sigma}{}^{\iota}{}_{\varkappa}\Lambda_{\lambda}{}^{\sigma}{}_{\mu} + \Delta_{\sigma}{}^{\iota}{}_{\lambda}\Lambda_{\mu}{}^{\sigma}{}_{\varkappa} + \Delta_{\sigma}{}^{\iota}{}_{\mu}\Lambda_{\varkappa}{}^{\sigma}{}_{\lambda}).$$

Diese Identität formen wir dadurch um, daß wir statt der gewöhnlichen Ableitungen des Tensors Λ die kovarianten Ableitungen einführen (gemäß (17)); wir erhalten so die Identität

$$0 \equiv (\Lambda_{\varkappa}{}^{\iota}{}_{\lambda;\mu} + \Lambda_{\lambda}{}^{\iota}{}_{\mu;\varkappa} + \Lambda_{\mu}{}^{\iota}{}_{\varkappa;\lambda}) + (\Lambda_{\varkappa}{}^{\iota}{}_{\alpha}\Lambda_{\lambda}{}^{\alpha}{}_{\mu} + \Lambda_{\lambda}{}^{\iota}{}_{\alpha}\Lambda_{\mu}{}^{\alpha}{}_{\varkappa} + \Lambda_{\mu}{}^{\iota}{}_{\alpha}\Lambda_{\varkappa}{}^{\alpha}{}_{\lambda}). \tag{24}$$

Sie ist die Bedingung dafür, daß sich die Λ in der angegebenen Weise durch die h ausdrücken lassen.

Durch Verjüngen dieser Gleichung nach den Indizes ι und μ erhält man ferner die Identität

$$0 \equiv \Lambda_{\varkappa}{}^{\alpha}{}_{\lambda;\alpha} + \varphi_{\lambda;\varkappa} - \varphi_{\varkappa;\lambda} - \varphi_{\alpha}\Lambda_{\varkappa}{}^{\alpha}{}_{\lambda}$$

oder

$$\Lambda_{\varkappa}{}^{\alpha}{}_{\lambda;\alpha} \equiv \varphi_{\varkappa,\lambda} - \varphi_{\lambda,\varkappa}, \tag{25}$$

wobei φ_λ die Abkürzung für $\Lambda_{\lambda}{}^{\alpha}{}_{\alpha}$ ist (22).

§ 4.

Die Feldgleichungen.

Die gesuchten einfachsten Feldgleichungen werden Bedingungen sein, denen der Tensor $\Lambda_{\mu}{}^{\alpha}{}_{\nu}$ zu unterwerfen ist. Da die Zahl der h-Komponenten n^2 ist und von diesen wegen der allgemeinen Kovarianz n unbestimmt bleiben müssen, so wird die Zahl der voneinander unabhängigen Feldgleichungen $n^2 - n$ sein müssen. Andererseits ist klar, daß eine Theorie desto befriedigender ist, je mehr sie die Möglichkeiten einschränkt (ohne mit Erfahrungen in Widerspruch zu treten). Die Zahl Z der Feldgleichungen soll also möglichst groß sein. Ist $\bar{Z}$ die Zahl der zwischen diesen bestehenden Identitäten, so muß $Z - \bar{Z}$ gleich $n^2 - n$ sein.

Gemäß der Vertauschungsregel der Differentiation ist

$$\Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu};\nu;\alpha} - \Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu};\alpha;\nu} - \Lambda_{\underline{\mu}}{}^{\sigma}{}_{\underline{\nu};\alpha}\Lambda_{\sigma}{}^{\alpha}{}_{\nu} \equiv 0. \tag{26}$$

Hierbei bedeuten die Striche unter einem Index das „Heraufziehen“ bzw. „Herunterziehen“ eines Index, also z. B.

$$\Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu}} \equiv \Lambda_{\beta}{}^{\alpha}{}_{\gamma} g^{\mu\beta} g^{\nu\gamma},$$

$$\Lambda_{\mu}{}^{\underline{\alpha}}{}_{\nu} \equiv \Lambda_{\mu}{}^{\beta}{}_{\nu} g_{\alpha\beta}.$$

Die Identität (26) schreiben wir nun in der Form

$$\text{(26a)} \qquad G^{\mu\alpha}{}_{;\alpha} - F^{\mu\nu}{}_{;\nu} + \Lambda_{\underline{\mu}}{}^{\sigma}{}_{\underline{\tau}} F_{\sigma\tau} \equiv 0,$$

wobei gesetzt ist

$$\text{(27)} \qquad G^{\mu\alpha} \equiv \Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu};\nu} - \Lambda_{\underline{\mu}}{}^{\sigma}{}_{\underline{\tau}} \Lambda_{\sigma}{}^{\alpha}{}_{\tau},$$

$$\text{(28)} \qquad F^{\mu\nu} \equiv \Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu};\alpha}.$$

Nun setzen wir als *Feldgleichungen* an:

$$\text{(29)} \qquad G^{\mu\alpha} = 0,$$

$$\text{(30)} \qquad F^{\mu\alpha} = 0.$$

Diese Gleichungen scheinen eine unerlaubte Überbestimmung zu enthalten. Denn ihre Anzahl ist $n^2 + \frac{n(n-1)}{2}$, während von ihnen vorläufig nur bekannt ist, daß sie den n Identitäten (26a) genügen.

Aus der Identität (25) in Verbindung mit (30) geht aber hervor, daß die $\varphi_\varkappa$ von einem Potential ableitbar sind. Wir setzen demgemäß

$$\text{(31)} \qquad F_\varkappa \equiv \varphi_\varkappa - \frac{\partial \lg \psi}{\partial x^\varkappa} = 0.$$

(31) ist mit (30) völlig äquivalent. Die Gleichungen (29), (31) sind zusammen $n^2 + n$ Gleichungen für die $n^2 + 1$ Funktionen $h_{s\nu}$ und ψ. Zwischen diesen Gleichungen besteht aber außer (26a) noch ein weiteres System von Identitäten, das wir nun ableiten wollen.

Bezeichnet $\underline{G}^{\mu\alpha}$ den antisymmetrischen Teil von $G^{\mu\alpha}$, so erhält man durch unmittelbares Ausrechnen aus (27)

$$\text{(32)} \qquad 2\underline{G}^{\mu\alpha} \equiv - S_{\mu}{}^{\nu}{}_{\alpha;\nu} + \tfrac{1}{2} S_{\underline{\sigma}}{}^{\mu}{}_{\underline{\tau}} \Lambda_{\sigma}{}^{\alpha}{}_{\tau} - \tfrac{1}{2} S_{\underline{\sigma}}{}^{\alpha}{}_{\underline{\tau}} \Lambda_{\sigma}{}^{\mu}{}_{\tau} + F^{\mu\alpha},$$

wobei zur Abkürzung der in allen Indizes antisymmetrische Tensor

$$\text{(33)} \qquad S_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu}} = \Lambda_{\underline{\mu}}{}^{\alpha}{}_{\underline{\nu}} + \Lambda_{\underline{\alpha}}{}^{\nu}{}_{\underline{\mu}} + \Lambda_{\underline{\nu}}{}^{\mu}{}_{\underline{\alpha}}$$

eingeführt ist. Durch Ausrechnen des ersten Gliedes von (32) ergibt sich

$$\text{(34)} \qquad 2\underline{G}^{\mu\alpha} \equiv - S_{\underline{\mu}}{}^{\nu}{}_{\underline{\alpha},\nu} - S_{\underline{\mu}}{}^{\sigma}{}_{\underline{\alpha}} \Lambda_{\sigma}{}^{\nu}{}_{\nu} + F^{\mu\alpha}.$$

Nun ist aber mit Rücksicht auf die Definition von F_k (31)

$$\Delta_{\sigma}{}^{\nu}{}_{\nu} - \Delta_{\nu}{}^{\nu}{}_{\sigma} \equiv \Lambda_{\sigma}{}^{\nu}{}_{\nu} \equiv \varphi_\sigma \equiv F_\sigma + \frac{\partial \lg \psi}{\partial x^\sigma}$$

oder

$$\Delta_{\sigma\ \nu}^{\ \nu} = \frac{\partial \lg \psi h}{\partial x^{\sigma}} + F_{\sigma}. \tag{35}$$

(34) nimmt daher die Form an

$$h\psi(2\underline{G}^{\mu\alpha} - F^{\mu\alpha} + S_{\underline{\mu}\ \underline{\alpha}}^{\ \sigma} F_{\sigma}) \equiv -\frac{\partial}{\partial x^{\sigma}}(h\psi S_{\underline{\mu}\ \underline{\alpha}}^{\ \sigma}). \tag{34b}$$

Wegen der Antisymmetrie folgt hieraus das gesuchte System identischer Gleichungen

$$\frac{\partial}{\partial x^{\alpha}}[h\psi(2\underline{G}^{\mu\alpha} - F^{\mu\alpha} + S_{\underline{\mu}\ \underline{\alpha}}^{\ \sigma} F_{\sigma})] \equiv 0. \tag{36}$$

Dies sind an sich n Identitäten, von welchen aber nur $n-1$ voneinander unabhängig sind, indem wegen der Antisymmetrie $[\ ]_{,\alpha,\mu} \equiv 0$ unabhängig davon gilt, was für $G^{\mu\alpha}$ und F_{μ} eingesetzt wird.

In den Identitäten (26a) und (36) ist $F^{\mu\alpha}$ durch F_{μ} ausgedrückt zu denken gemäß der aus (31) folgenden Beziehung

$$F_{\mu\alpha} \equiv F_{\mu,\alpha} - F_{\alpha,\mu}. \tag{31a}$$

Wir sind nun in der Lage, die Kompatibilität der Feldgleichungen (29), (30) bzw. (29), (31) zu beweisen.

Zunächst ist zu zeigen, daß die um die Zahl der (unabhängigen) Identitäten verminderte Zahl der Feldgleichungen um n kleiner ist als die Zahl der Feldvariablen. Wir haben

Zahl der Gleichungen (29), (31):	$n^2 + n$,
Zahl der (unabhängigen) Identitäten:	$n + n - 1$,
Zahl der Feldvariablen:	$n^2 + 1$,

$$(n^2 + n) - (n + n - 1) = (n^2 + 1) - n.$$

Die Zahl der Identitäten ist also gerade die richtige. Damit begnügen wir uns aber nicht, sondern beweisen folgenden

Satz. *Sind in einem Schnitt $x^n =$ konst. alle Differentialgleichungen erfüllt und außerdem* **überall** *$(n^2+1) - n$ derselben (passend ausgewählt), so sind alle $n^2 + n$ Gleichungen von selbst überall erfüllt.*

Beweis. Es seien alle Gleichungen im Schnitte $x^n = a$ erfüllt, außerdem überall die Gleichungen, welche dem Nullsetzen von

$$\begin{array}{lll} F_1 & \ldots F_{n-1} & F_n \\ G^{11} & \ldots G^{1\,n-1} & \\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \\ G^{n-1\,1} & \ldots G^{n-1\,n-1} & \end{array}$$

entsprechen. Aus (31a) folgt zunächst, daß dann die $F^{\mu\alpha}$ überall verschwinden. Nun folgert man aus (36), daß in dem benachbarten Schnitte

$x^n = a + da$ auch die antisymmetrischen $\underline{G}^{\mu\alpha}$ für $\alpha = n$ verschwinden müssen[1]). Ferner folgt dann analog aus (26a), daß außerdem noch die symmetrischen $\underline{\underline{G}}^{\mu\alpha}$ für $\alpha = n$ für den Nachbarschnitt $x^n = a + da$ verschwinden müssen. Durch Wiederholung dieser Schlußweise folgt die Behauptung.

§ 5.

Erste Näherung.

Wir betrachten nun ein Feld, welches sich von einem euklidischen mit gewöhnlichem Parallelismus nur unendlich wenig unterscheidet. Dann können wir setzen

(37) $$h_{s\nu} = \delta_{s\nu} + \overline{h}_{s\nu},$$

wobei die $\overline{h}_{s\nu}$ unendlich klein erster Ordnung sind und kleine Größen höherer Ordnung vernachlässigt werden. Dann ist gemäß (5) bzw. (6) zu setzen

(38) $$h_s^{\ \nu} = \delta_{s\nu} - \overline{h}_{\nu s}.$$

Die Feldgleichungen (29), (30) lauten in erster Näherung

(39) $$\overline{h}_{a\mu,\nu,\nu} - \overline{h}_{a\nu,\nu,\mu} = 0,$$

(40) $$\overline{h}_{a\mu,a,\nu} - \overline{h}_{a\nu,a,\mu} = 0.$$

Gleichung (40) ersetzen wir durch

(40a) $$\overline{h}_{a\nu,a} = \chi_{,\nu}.$$

Wir behaupten nun, daß es eine infinitesimale Koordinatentransformation $x^{\nu\prime} = x^\nu - \xi^\nu$ gibt, welche die Größen $\overline{h}_{a\nu,\nu}$ und $\overline{h}_{a\nu,a}$ sämtlich zum Verschwinden bringt.

Beweis. Man beweist zunächst, daß

(41) $$\overline{h}'_{\mu\nu} = \overline{h}_{\mu\nu} + \xi^\mu{}_{,\nu}.$$

Hieraus

$$\overline{h}'_{a\nu,\nu} = h_{a\nu,\nu} + \xi^a{}_{,\nu,\nu},$$
$$\overline{h}'_{a\nu a} = \overline{h}_{a\nu a} + \xi^a{}_{,a,\nu}.$$

Die rechten Seiten verschwinden mit Rücksicht auf (40a), wenn die Gleichungen erfüllt sind

(42) $$\xi^a{}_{,\nu,\nu} = -\overline{h}_{a\nu,\nu},$$
$$\xi^a{}_{,a} = -\chi.$$

[1]) Es verschwinden für $x^n = a$ die $\dfrac{\partial \underline{G}^{\mu n}}{\partial x^n}$.

Diese $n+1$ Gleichungen für die n Größen ξ^{a} sind aber kompatibel, weil gemäß (40a)

$$(-\bar{h}_{a\nu,\nu})_{,a} - (-\chi)_{,\nu,\nu} = 0.$$

Bei der neuen Koordinatenwahl lauten die Feldgleichungen

$$\bar{h}_{a\mu,\nu\nu} = 0,$$
$$\bar{h}_{a\mu,a} = 0,$$
$$\bar{h}_{a\mu,\mu} = 0.$$

Spalten wir nun die $\bar{h}_{a\mu}$ gemäß den Gleichungen

$$\bar{h}_{a\mu} + \bar{h}_{\mu a} = \bar{g}_{a\mu},$$
$$\bar{h}_{a\mu} - \bar{h}_{\mu a} = a_{a\mu},$$

wobei $\delta_{a\mu} + \bar{g}_{a\mu}$ $(= g_{\mu\nu})$ in erster Näherung die Metrik bestimmen, so nehmen die Feldgleichungen die übersichtliche Form an

$$\bar{g}_{a\mu,\sigma,\sigma} = 0, \tag{44}$$
$$\bar{g}_{a\mu,\mu} = 0, \tag{45}$$
$$a_{a\mu,\sigma,\sigma} = 0, \tag{46}$$
$$a_{a\mu,\mu} = 0. \tag{47}$$

Die Auffassung liegt nahe, daß die $\bar{g}_{a\mu}$ das Gravitationsfeld, die $a_{a\mu}$ das elektromagnettsche Feld in erster Näherung darstellen. (44), (45) entsprechen der Poissonschen Gleichung, (46), (47) den Maxwellschen Gleichungen des leeren Raumes. Es ist interessant, daß die Feldgesetze der Gravitation von denen des elektromagnetischen Feldes separiert erscheinen, wie es der Erfahrung von der Unabhängigkeit beider Felder entspricht. In Strenge kommt aber nach dieser Theorie keinem dieser Felder eine Sonderexistenz zu.

Bezüglich der Kovarianz der Gleichungen (44) bis (47) gilt folgendes. Für die $h_{s\mu}$ gilt allgemein das Transformationsgesetz

$$h'_{s\mu} = \alpha_{st} \frac{\partial x^{\sigma}}{\partial x^{\mu\prime}} h_{t\sigma}.$$

Wählt man die Koordinatentransformation linear und orthogonal sowie konform der Drehung der Lokalsysteme, also

$$x^{\mu\prime} = \alpha_{\mu\sigma} x^{\sigma}, \tag{48}$$

so ergibt sich als Transformationsgesetz

$$h'_{s\mu} = \alpha_{st}\alpha_{\mu\sigma} h_{t\sigma}, \tag{49}$$

also genau dasselbe wie für Tensoren in der speziellen Relativitätstheorie. Da dasselbe Transformationsgesetz wegen (48) für die $\delta_{s\mu}$ gilt, so gilt es auch für die Größen $\bar{h}_{a\mu}$, $\bar{g}_{a\mu}$ und $a_{a\mu}$. Bezüglich solcher Transformationen sind die Gleichungen (44) bis (47) kovariant.

Schlußbemerkung.

Der große Reiz der hier dargelegten Theorie liegt für mich in ihrer Einheitlichkeit und in der hochgradigen (erlaubten) Überbestimmung der Feldvariablen. Auch habe ich zeigen können, daß die Feldgleichungen in erster Näherung auf Gleichungen führen, welche der Newton-Poissonschen Theorie der Gravitation sowie der Maxwellschen Theorie des elektromagnetischen Feldes entsprechen. Trotzdem bin ich noch weit davon entfernt, die physikalische Gültigkeit der abgeleiteten Gleichungen behaupten zu können. Der Grund liegt darin, daß mir die Ableitung von Bewegungsgesetzen für die Korpuskeln noch nicht gelungen ist.

(Eingegangen am 19. 8. 1929.)

Notice historique sur la notion de parallélisme absolu.

Von

E. Cartan in Paris.

M. Einstein, à qui j'avais signalé certains de mes travaux contenant la notion de variété riemannienne à parallélisme absolu, a bien voulu me demander d'écrire une notice historique sur cette notion, envisagée du point de vue géométrique. Je le fais d'autant plus volontiers qu'à côté des questions de priorité, qui n'intéressent après tout qu'un petit nombre de personnes, il existe plusieurs problèmes que j'aurai ainsi l'occasion de signaler et dont la solution est susceptible d'intéresser les physiciens. Je m'étendrai de préférence sur l'aspect géométrique des questions, laissant à l'arrière-plan les développements analytiques correspondants.

I.

1. La notion de parallélisme absolu (Fernparallelismus) dans une variété riemannienne peut être conçue indépendamment de toute idée métrique. Supposons la variété à n dimensions. Deux vecteurs infiniment petits d'origines différentes seront dits *parallèles* (ou plutôt *équipollents*) si, pour ces deux vecteurs, n formes de Pfaff linéairement indépendantes

$$^{i}L = {}^{i}h_k\, dx^k \qquad (i = 1, 2, \ldots, n)$$

sont numériquement égales chacune à chacune. On a naturellement le même parallélisme absolu si l'on substitue aux n formes ^{i}L n combinaisons linéaires indépendantes à coefficients *constants* de ces formes.

M. Weitzenböck a en 1923 (*4* [1]), p. 320), et même en 1921 (*1*, p. 51), défini une certaine dérivation covariante par rapport à un système de n formes de Pfaff linéairement indépendantes. Mais on ne peut voir dans cette opération purement formelle la première apparition de la notion

[1]) Les nombres en italique renvoient aux articles cités dans l'index bibliographique, placé à la fin de la notice.

de parallélisme absolu. Ricci, dans sa méthode de calcul sur les n-uples de congruences orthogonales, qui remonte à 1895, utilise au fond, dans l'étude des variétés riemanniennes, un système de n expressions de Pfaff; c'est aussi ce qu'on fait toutes les fois qu'en Géométrie différentielle on se sert de systèmes de référence mobiles *locaux*. Il y a là une méthode générale qui est en soi tout à fait étrangère à la notion de parallélisme absolu [2]).

2. Cette notion est au contraire explicitement introduite en 1923 dans un mémoire (*5a*) consacré au développement d'une théorie générale, que j'avais esquissée l'année précédente dans deux notes des *Comptes Rendus* (*2* et *3*), et que j'ai exposée ensuite sous ses divers aspects géométriques dans plusieurs articles ou conférences (*6*, *9*, *19*).

Cette théorie fait correspondre à chaque espace à groupe fondamental, au sens de F. Klein (espace euclidien, espace affine, espace projectif, etc.), des espaces *non holonomes*, également à groupe fondamental (espace à connexion euclidienne, affine, projective, etc.). Les espaces de Riemann, tels qu'on les conçoit dans la théorie classique, rentrent dans la classe plus générale des espaces à connexion euclidienne, dont le groupe fondamental est le groupe des déplacements euclidiens.

Un espace général à connexion euclidienne peut être conçu comme formé d'une infinité de morceaux infiniment petits d'espace euclidien, avec une loi de raccord permettant d'intégrer deux morceaux contigus dans un seul et même espace euclidien. Voici, d'une manière plus précise, la nature de cette loi de raccord. Considérons deux points infiniment voisins A et A', ainsi que deux systèmes de référence rectangulaires locaux (R_A) et $(R_{A'})$ attachés à ces points. Un observateur placé en A pourra s'imaginer être dans un espace euclidien, et il saura, la loi de raccord étant connue, localiser dans cet espace euclidien le point A' et le repère $(R_{A'})$; autrement dit, il connaîtra les coordonnées rectangulaires de A' par rapport à (R_A), ce qui revient à connaître le ds^2 de l'espace, et les angles que font les axes de $(R_{A'})$ avec ceux de (R_A), ce qui revient à connaître la loi de transport par parallélisme. Il connaîtra par suite l'angle d'un vecteur quelconque issu de A' avec un vecteur quelconque issu de A. Si l'on imagine une suite continue d'observateurs échelonnés le long d'un arc de courbe AB, l'observateur placé en A sera donc capable de localiser de proche en proche dans un même espace euclidien (espace

[2]) Cela ne veut pas dire que les recherches de M. Weitzenböck n'aient aucune portée géométrique, car elles sont immédiatement utilisables dans la théorie analytique du parallélisme absolu, une fois cette notion géométrique acquise. Voir les mémoires *4*, *10*, *11*, *12*, *15*, *21*, *23* qui contiennent les développements analytiques de la théorie de M. Weitzenböck. Cf. note [3]) du n° 7.

euclidien tangent en A) les différents points de AB et les différents vecteurs issus de ces points; on pourra dire qu'il saura *développer* sur son espace euclidien la ligne AB et la portion de l'espace qui avoisine immédiatement cette ligne.

L'observateur A sera averti qu'il n'est pas dans un vrai espace euclidien s'il essaie, en suivant deux chemins différents ACB et $AC'B$, de localiser dans son espace euclidien le point B et les vecteurs issus de B. Suivant le chemin suivi, il n'attribuera pas, *dans son espace euclidien*, la même position au point B, pas plus qu'il n'attribuera aux vecteurs issus de B la même orientation. La rotation que lui paraissent avoir subie les vecteurs en passant d'un chemin à l'autre constitue la *courbure* associée au cycle $BC'ACB$; la translation qui amène en coïncidence les deux positions différentes attribuées au point B constitue la *torsion* associée à ce même cycle; le vecteur qui répresente cette translation est le *vecteur de torsion* du cycle. Si le cycle est infiniment petit, la courbure se traduit analytiquement par le tenseur bien connu à quatre indices, la torsion par le tenseur à trois indices A_{ij}^{k} utilisé par M. Einstein. La condition nécessaire et suffisante pour que le tenseur de torsion soit nul est que le transport par parallélisme soit celui qui a été défini en 1917 par M. Levi-Civita (*geodätische Übertragung* de M. Schouten).

Tout ce qui précède s'étend, *mutatis mutandis*, aux espaces à connexion affine.

3. Revenons maintenant au parallélisme absolu. J'ai démontré (*5a*, p. 368), ce qui n'est pas absolument évident, que si la courbure associée à tout cycle *infiniment petit* est nulle (espace sans courbure), l'espace est doué d'un parallélisme absolu; autrement dit, un vecteur issu d'un point A, transporté parallèlement à lui-même de proche en proche de A en B, donne toujours le même vecteur final (pourvu toutefois que les chemins intermédiaires suivis soient réductibles les uns aux autres par déformation continue). Si l'on choisit en un point A un système de référence formé de n vecteurs indépendants et si l'on prend en un point quelconque M le système de référence formé des n vecteurs parallèles aux premiers, la connexion affine de l'espace est complètement définie (*5a*, p. 368; *5c*, p. 20) par les n formes de Pfaff ω^i qui représentent les projections d'un vecteur infiniment petit sur les axes de coordonnées locaux attachés à l'origine du vecteur[3]).

---

[3]) Le vecteur de torsion associé à un cycle ne peut se définir avec précision que si l'on choisit l'origine du cycle, à moins que le cycle ne soit infiniment petit. Il en est autrement s'il y a un parallélisme absolu (*16*, p. 37). Avec les notations du n° 1, le vecteur de torsion associé à un cycle fini est celui qui a pour composantes,

La démonstration, qui est donnée dans le cas général d'un espace à connexion affine, est naturellement valable dans le cas particulier d'un espace à connexion euclidienne; on obtient alors les espaces riemanniens à parallélisme absolu de M. Einstein. J'ai du reste signalé, toujours dans le même mémoire (*5a*, p. 404—409; cf. *6*, p. 301—302) l'exemple le plus simple d'un tel espace: c'est, pour $n=2$, celui de la surface terrestre, supposée sphérique, où l'on regarde comme parallèles deux directions qui font le même angle avec l'aiguille aimantée; le vecteur de torsion est ici tangent aux cercles méridiens.

Il est intéressant de remarquer que la première théorie de la relativité de M. Einstein repose sur la notion d'espace riemannien sans torsion, alors que la théorie actuelle repose sur celle d'espace riemannien sans courbure.

4. Faisons ici la remarque évidente qu'on peut passer d'un espace à connexion affine sans courbure à un espace riemannien à parallélisme absolu en prenant comme forme différentielle quadratique fondamentale la somme des carrés des n expressions de Pfaff ω^i, les n vecteurs de coordonnées devenant unitaires et rectangulaires, — ou encore une forme quadratique à coefficients *constants* arbitraires construite avec les ω^i, les n vecteurs de coordonnées formant une figure invariable de grandeur et de forme. Inversement on peut arriver au parallélisme absolu le plus général d'un espace de Riemann donné en décomposant son ds^2 en une somme de n carrés.

5. Dans le cas d'un espace à connexion *affine*, le tenseur de torsion Λ^k_{ij} se décompose (*5c*, p. 30—33) en deux tenseurs *irréductibles*. L'un est le vecteur $\Lambda^k_{ik}=\varphi_i$ de M. Einstein, qui a une signification purement affine. L'autre peut être interprété géométriquement: il est nul dans le cas et dans le cas seulement où le vecteur de torsion associé à un cycle élémentaire est situé dans l'élément plan de ce cycle; les espaces correspondants sont les espaces à connexion *semi-symétrique* de M. J. A. Schouten [4]).

---

par rapport aux vecteurs de référence choisis, les n intégrales $\int {}'h_k\,dx^k$ étendues au cycle. Le théorème général de la conservation de la courbure et de la torsion (*5a*, p. 373—375), qui comprend en particulier les identités de Bianchi, revient ici à un théorème classique de H. Poincaré (Acta Math. **9** (1887), p. 321); géométriquement, il signifie que la somme géométrique des vecteurs de torsion associés aux éléments d'une surface fermée est nulle.

[4]) Les connexions affines que j'ai introduites rentrent dans des connexions encore plus générales dues à M. Schouten (Math. Zeitschr. **13** (1922), p. 56—81); mais le point de vue de M. Schouten est différent du mien. Pour lui le transport parallèle (*lineare Übertragung*) est la notion géométrique essentielle; pour moi, elle n'est qu'un moyen qui tient aux propriétés particulières de l'espace affine et qui ne peut plus s'utiliser, au moins directement, pour établir la notion d'espace à connexion projective (ou conforme, etc.).

Si l'espace est à connexion *euclidienne*, le second tenseur de torsion cesse d'être irréductible (*5c*, p. 50—52); en particulier dans le cas, important pour la relativité, où n est égal à 4, l'un des deux tenseurs irréductibles dans lesquels il se décompose est un *vecteur* ψ_i, *qui a ainsi une signification essentiellement métrique* (*5c*, p. 69—71). Avec les notations ordinaires, on a

$$\psi_i = \frac{1}{\sqrt{-g}}\left(g_{j\alpha}\Lambda^{\alpha}_{kh} + g_{k\alpha}\Lambda^{\alpha}_{hj} + g_{h\alpha}\Lambda^{\alpha}_{jk}\right),$$

les indices i, j, k, h formant une permutation *paire* des indices 1, 2, 3, 4.

6. La dérivation covariante par rapport à un système de n expressions de Pfaff a été, après M. Weitzenböck, découverte de nouveau par M. G. Vitali (*7* et *8*) en 1924. Mais cet auteur lui attache une signification géométrique et reconnaît la possibilité d'en déduire une connexion affine, qu'il démontre être sans courbure. Le théorème réciproque, que j'avais démontré en 1923, a été démontré de nouveau par M. E. Bortolotti en 1927 dans le cas d'une connexion euclidienne (*17*). Depuis, le parallélisme absolu a été considéré par différents auteurs, dont on trouvera la liste, probablement incomplète, dans l'index bibliographique.

7. Je voudrais maintenant donner un aperçu rapide des principaux problèmes qu'on s'est posés relativement au parallélisme absolu.

Plaçons-nous d'abord au point de vue strictement *affine*. Nous avons montré, M. Schouten et moi (*13*), en 1926, que dans l'espace représentatif des transformations d'un groupe fini et continu, il existait deux parallélismes absolus remarquables. Si l'on désigne pour abréger par T_x la transformation générale du groupe, de paramètres $x_1, x_2, \ldots, x_n$, les expressions ω^i qui définissent le premier parallélisme absolu sont les paramètres de la transformation infinitésimale $T_x^{-1} T_{x+dx}$; celles qui définissent le second parallélisme sont les paramètres de la transformation infinitésimale $T_{x+dx} T_x^{-1}$.[5]) Les vecteurs de torsion correspondant à ces deux parallélismes sont égaux et opposés, et les quantités Λ^k_{ij} ne sont, au signe près, que les *constantes de structure* c_{ijk} de S. Lie.

---

[5]) Les expressions de Pfaff ω^i jouent un rôle important dans ma théorie de la structure des groupes continus, théorie d'où découle une méthode générale de Géométrie différentielle par l'utilisation d'un système de référence mobile. J'ai d'autre part [Ann. Ec. Norm. 25 (1908), p. 60—83] ramené la recherche des invariants différentiels d'un système différentiel quelconque vis-à-vis d'un groupe de transformations continu, *fini ou infini*, à la recherche des invariants d'un système de n expressions de Pfaff indépendantes à n variables vis-à-vis du groupe général de ces n variables. La solution n'exige, comme opérations analytiques, que la dérivation covariante d'un scalaire par rapport au système donné d'expressions de Pfaff et la formation du *covariant bilinéaire* (rotationnel) d'une expression de Pfaff.

8. Ces espaces de groupes présentent un intérêt physique. En effet, avec la nouvelle théorie de M. Einstein, il est naturel d'appeler *homogène* un Univers où les vecteurs de torsion associés à deux éléments de surface *parallèles* sont eux-mêmes *parallèles*, c'est à dire où le transport parallèle conserve la torsion. Or (*13*, p. 813; *16*, p. 50—51) les seuls espaces à parallélisme absolu jouissant de cette propriété sont les espaces représentatifs des groupes.

On peut encore les caractériser autrement (*16*). Appelons *translation infinitésimale*, dans un espace à parallélisme absolu, une transformation ponctuelle par laquelle les différents points de l'espace décrivent des vecteurs infiniment petits équipollents. On peut associer à la connexion affine sans courbure définie par le parallélisme absolu donné une seconde connexion affine, comportant en général courbure et torsion; il suffit (*16*, p. 52—53) de convenir que deux vecteurs d'origines infiniment voisines sont parallèles (au second sens) s'ils se déduisent l'un de l'autre par la translation infinitésimale qui amène en coïncidence leurs deux origines· La torsion de cette nouvelle connexion est toujours égale et opposée à celle de la première. Pour que la nouvelle connexion soit, elle aussi, sans courbure, il faut et il suffit que l'espace donné soit un espace de groupe (*16*, p. 53): les deux parallélismes absolus de cet espace se déduisent alors l'un de l'autre par le procédé qui vient d'être indiqué.

9. Quel que soit le ds^2 qu'on attribue à un espace de groupe pour en faire un espace riemannien *homogène* à parallélisme absolu, le vecteur φ_i est toujours le même, *et il se trouve que son rotationnel est toujours nul, ce qui exclurait donc l'Electromagnétisme de tout Univers homogène.* Cette conclusion serait en défaut si l'on pouvait définir le potentiel électromagnétique au moyen du vecteur ψ_i (n° 5); mais nous sortons là du domaine de la Géométrie. Remarquons simplement qu'en principe les phénomènes mécaniques sont de nature purement affine, tandis que les phénomènes életromagnétiques sont de nature essentiellement métrique; il peut donc paraître assez naturel de chercher à représenter le potentiel électromagnétique par un vecteur non purement affine.

10. Un autre problème dont nous nous sommes également occupés, M. Schouten et moi, en 1926 (*14*) se rapporte exclusivement aux espaces riemanniens à parallélisme absolu. Est-il possible, dans un espace riemannien donné par son ds^2, de définir un parallélisme absolu tel que les géodésiques de ce parallélisme se confondent avec les géodésiques riemanniennes? On peut formuler ce problème de bien d'autres manières. Par exemple on peut, en s'appuyant sur un théorème général que j'ai démontré en 1923 (*5a*, p. 408), se demander s'il est possible de trouver un parallélisme

absolu tel que le vecteur de torsion associé à un élément de surface quelconque soit normal à cet élément. On peut encore se demander dans quels cas la connexion affine associée au parallélisme absolu suivant le procédé indiqué au n° 8 conserve les longueurs des vecteurs. Enfin on peut rattacher la question à un problème de Mécanique classique: est-il possible, étant donné un système matériel à n degrés de liberté, de choisir des *caractéristiques des vitesses* p_i telles que les mouvements spontanés du système soient donnés par les équations $\frac{dp_i}{dt} = 0$ *); c'est ce qui se présente par exemple si l'on considère un corps solide mobile autour d'un point fixe O, l'ellipsoïde d'inertie relatif à O étant une sphère, et si l'on prend pour caractéristiques des vitesses les composantes p, q, r de la rotation instantanée autour de O.

11. Nous avons réussi à résoudre complètement le problème, du moins dans le cas où le ds^2 donné est *défini*. Si l'on se borne aux solutions *irréductibles*, dont toutes les autres se déduisent facilement, on trouve

1° les espaces représentatifs des groupes simples clos, doués d'un ds^2 intrinsèquement lié à la structure du groupe, le parallélisme absolu étant l'un quelconque des deux parallélismes absolus attachés au groupe;

2° l'espace elliptique à 7 dimensions, qui admet deux familles *continues* de parallélismes absolus satisfaisant aux conditions voulues; une étude de ces parallélismes a été faite par M. Vaney (*26*).

En particulier l'espace elliptique (ou l'espace sphérique) à trois dimensions rentre dans la première catégorie: les deux parallélismes absolus dont il est question plus haut ont été signalés depuis longtemps par Clifford. Cet espace est l'espace représentatif du groupe des rotations de l'espace ordinaire; au point de vue mécanique, ses différents points représentent les différentes positions d'un corps solide mobile autour d'un point fixe; les deux parallélismes admettent alors une interprétation cinématique remarquable (*6*, p. 305—308). On voit que les parallélismes de Clifford, qui formaient un chapitre tout à fait isolé de la Géométrie, sont maintenant rattachés à une théorie très générale comprenant, malgré l'opposition apparente des deux notions, le parallélisme de M. Levi-Civita et les parallélismes de Clifford.

12. Les espaces riemanniens dont il vient d'être question rentrent dans une catégorie plus générale, celle des espaces dans lesquels le transport parallèle conserve la courbure et la torsion; ils admettent alors un

*) Ce problème de Mécanique a fait l'objet des recherches de M. Georg Hamel [Zeitschr. für Math. u. Phys. **50** (1904), p. 1—53], qui avait trouvé une partie des solutions indiquées ci-dessous (n° 11).

groupe transitif de déplacements rigides laissant également invariantes la courbure et la torsion. Réciproquement si un espace riemannien, envisagé du point de vue classique, admet un groupe transitif de déplacements rigides, c'est-à-dire laissant invariant le ds^2, on peut toujours (du moins si le ds^2 est défini), définir dans cet espace une connexion euclidienne telle que le transport parallèle correspondant conserve la courbure et la torsion. Le vecteur φ_i a, là encore, son rotationnel toujours nul. Il est vrai que, dans les applications possibles à la théorie de la relativité, le ds^2 est indéfini; mais même dans ce cas, pour $n=4$, la conclusion subsiste. Les espaces sans torsion dans lesquels le transport parallèle conserve la courbure jouent un rôle important en Géométrie, mais ils sortent tout à fait du cadre de cette notice [7]).

Index bibliographique.

1. R. Weitzenböck, Neuere Arbeiten der algebraischen Invariantentheorie, Differentialinvarianten, Enzyklopädie **3** (3), Heft 6, 1922.

2. E. Cartan, Sur une généralisation de la notion de courbure de Riemann et les espaces à torsion, Comptes rendus **174** (1922), p. 593—595.

3. E. Cartan, Sur les espaces généralisés et la théorie de la relativité, id. **174** (1922), p. 734—737.

4. R. Weitzenböck, Invariantentheorie (Groningen, Noordhoff, 1923).

5. E. Cartan, Sur les variétés à connexion affine et la théorie de la relativité généralisée, *a*) Annales Ec. Norm. **40** (1923), p. 325—412; *b*) id. **41** (1924), p. 1—25; *c*) id. **42** (1925), p. 17—88.

6. E. Cartan, Les récentes généralisations de la notion d'espace, Bull. Sc. Math. **48** (1924), p. 294—320.

7. G. Vitali, Una derivazione covariante formata coll' ausilio di *n* sistemi covarianti del 1° ordine, Atti Soc. Ligustica Sc. & Lett. **2** (1924), p. 248—253.

8. G. Vitali, Intorno ad una derivazione covariante nel calcolo assoluto, id. **4** (1925), p. 287—291.

9. E. Cartan, La théorie des groupes et les recherches récentes de Géométrie différentielle, L'Enseignement math. 1924/25, p. 1—18; Revista mat. hispano-americana 1927, p. 1—13; Proc. Int. Math. Congress Toronto 1924, **1** (1928), p. 85—94.

10. G. F. C. Griss, Differentialinvarianten von Systemen von Vektoren, Dissert., Amsterdam 1925.

11. R. Weitzenböck, Über Differentialinvarianten von kovarianten Tensoren, Proc. Akad. Amsterdam **29** (1926), p. 400—403.

12. R. Weitzenböck, Über Differentialinvarianten eines speziellen schiefsymmetrischen Tensors p_{ik} im R_4, id. **29** (1926), p. 404—409.

13. E. Cartan and J. A. Schouten, On the Geometry of the Group-manifold of simple and semi-simple groups, id. **29** (1926), p. 803—815.

---

[7]) On pourra consulter, sur d'autres problèmes de géométrie qu'on peut rattacher au parallélisme absolu, une note toute récente de M. E. Bortolotti (*25*).

14. E. Cartan and J. A. Schouten, On Riemannian Geometries admitting an absolute parallelism, id. **29** (1926), p. 933—946.

15. M. Euwe, Differentiaalinvarianten van twee covariante vectorvelden met vier veranderlijken, Dissert., Amsterdam 1926.

16. E. Cartan, La Géométrie des groupes de transformations, Journal Math. pures et appl. **6** (1927), p. 1—119.

17. E. Bortolotti, Parallelismi assoluti nelle V_n riemanniane, Atti Istit. Veneto **86** (1926/27), p. 456—465.

18. E. Bortolotti, On metric connexions with absolute parallelism, Proc. Akad. Amsterdam **30** (1927), p. 216—218.

19. E. Cartan, La théorie des groupes et la Géométrie, L'Enseignement math. **26** (1927), p. 200—225.

20. L. P. Eisenhart, Non-Riemannian Geometry, Amer. Math. Soc. Colloq. Publications VIII, New-York 1927.

21. W. Euwe, Differentiaalinvarianten en partieele differentiaalvergelijkingen uit de tensorrekening, Dissert., Amsterdam 1927.

22. A. Einstein, Riemann-Geometrie mit Aufrechterhaltung des Fernparallelismus, Sitzungsber. Berlin 1928, XVII, p. 217—221.

23. R. Weitzenböck, Differentialinvarianten in der Einsteinschen Theorie des Fernparallelismus, id. 1928, XXVI.

24. E. Bortolotti, Parallelismo assoluto nelle varietà a connessione affine, e nuove vedute sulla relatività, Accad. Sc. Bologna, 27 gennaio 1929.

25. E. Bortolotti, Stelle di congruenze e parallelismo assoluto, basi geometriche di una recente teoria di Einstein, Atti R. Accad. Lincei **9** (1929), p. 530—538.

26. F. Vaney, Le parallélisme absolu dans les espaces elliptiques réels à 3 et 7 dimensions et le principe de trialité dans l'espace elliptique à 7 dimensions (Thèse; Paris, Gauthier-Villars, 1929).

(Eingegangen am 19. 8. 1929.)

Über Weierstraßsche Approximation, besonders durch Hermitesche Interpolation*).

Von

Leopold Fejér in Budapest.

Einleitung.

1. Für den Weierstraßschen Approximationssatz, nach welchem man eine beliebige, im Intervalle $-1 \leqq x \leqq +1$ gegebene stetige Funktion in diesem Intervalle durch ein rationales ganzes Polynom $g(x)$ mit beliebig großer Genauigkeit gleichmäßig approximieren kann, existieren viele Beweise. Sie sind im Pringsheim-Molkschen[1]) Enzyklopädieartikel zusammengestellt, und die verschiedenen Vorzüge der verschiedenen Beweise sind dort treffend hervorgehoben. Von den älteren Beweisen wäre noch der von de la Vallée Poussin, von den neueren der von Dunham Jackson und Serge Bernstein hinzuzufügen.

2. Auch ich habe im Laufe der Zeit drei verschiedene Beweise für den Weierstraßschen Satz gefunden, die ich in den Jahren 1900, 1908 und 1915 veröffentlicht habe. Sie stützen sich unmittelbar auf die Fourier-Tschebyscheffsche, oder auf die Legendresche Reihe, oder endlich auf eine gewisse *Interpolationsreihe* der gegebenen Funktion. Meinen *Ausgangspunkt* bilden also die klassischen *Reihenentwicklungen* der willkürlichen Funktion, und *nicht* gewisse, übrigens wichtige, in der Analysis oder in der theoretischen Physik auftretende s. g. *singuläre Integrale*, oder anders geartete Kunstgriffe. Weiter gebe auch ich immer einfach lautende Herstellungsregeln für die gesuchten Näherungspolynome.

Die *erste* Methode ist mit der *zweiten* — wenn auch jede der beiden ihre interessante Besonderheit hat — eng verwandt; ist doch sowohl die Fouriersche Kosinusreihe als die Legendresche Reihe ein hervorragender

---

*) Vorgetragen an der Schlesischen Friedrich Wilhelms-Universität zu Breslau am 2. Juli 1929.

[1]) A. Pringsheim, J. Molk, Principes fondamentaux de la théorie des fonctions, Encyclopédie des sciences mathématiques pures et appliquées, tome 2, vol. 1; insbesondere Fußnote [188]).

Spezialfall der allgemeinen ultrasphärischen Entwicklung der willkürlichen Funktion, d. h. der Entwicklung nach s. g. ultrasphärischen Polynomen der unabhängigen Veränderlichen x.

Um nun diese beiden, in Rede stehenden Methoden in diesen einleitenden Zeilen ganz kurz zu charakterisieren, bemerke ich, daß nicht die *Partialsummen* selbst dieser Reihenentwicklungen, sondern ihre *arithmetischen Mittel* es sind, die jeweils bei einer beliebigen stetigen Funktion zum Ziele führen. Übrigens werde ich über diese Reihenentwicklungsmethode im folgenden nur skizzenhaft einige Worte sagen, und auch dies nur aus dem Grunde, um gewisse Zusammenhänge hervortreten zu lassen. Ausführlicher möchte ich aber von der *dritten* Methode, von der *Interpolationsmethode* sprechen, um so mehr, als ich in diesem Gegenstande in allerletzter Zeit etwas weiter gekommen bin. Hier dienen als Interpolationsabszissen — um es kurz auszudrücken — die Nullstellen der ultrasphärischen Polynome; also insbesondere die s. g. Fourier-Tschebyscheffschen Abszissen, oder die s. g. Legendre-Gaußschen Abszissen. Um auch mein diesbezügliches Resultat hier kurz zu charakterisieren, bemerke ich, daß, wenn eine beliebige stetige Funktion im Intervalle $(-1, +1)$ vorgeschrieben ist, *nicht die gewöhnlichen Lagrangeschen Interpolationspolynome, sondern gewisse höhere, s. g. Hermitesche Interpolationspolynome es sind, die zur Funktion gleichmäßig konvergieren.* Dabei verstehe ich unter Lagrangeschem Interpolationspolynom, wie üblich, dasjenige Polynom von höchstens $\overline{n-1}$-tem Grade, das an den vorgeschriebenen Stellen $x_1, x_2, \ldots, x_n$ denselben Wert hat wie die gegebene Funktion, unter Hermiteschem Interpolationspolynom ein Polynom, welches ebenfalls der Bedingung unterworfen ist, daß sein Wert an den Stellen $x_1, x_2, \ldots, x_n$ mit dem Funktionswert übereinstimmt, dessen Grad aber *höher als* $(n-1)$ *sein kann.*

Diese Bemerkung zeigt aber auch, daß die dritte Methode mit den zwei ersten verwandt ist; sind es doch die *ultrasphärischen Abszissengruppen* — wie ich sie nennen möchte —, die (hauptsächlich) als passende Interpolationsstellen dienen. Dabei entsprechen die höheren, Hermiteschen Interpolationsgebilde genau den Mittelgebilden bei den erwähnten Reihenentwicklungen.

§ 1.

Approximation einer stetigen Funktion durch eine rationale ganze Funktion mit Hilfe der Methode der kleinsten Quadrate, d. h. mit Benutzung der Teilsummen der Legendreschen Reihe der zu approximierenden stetigen Funktion.

3. Es sei $g(x)$ eine beliebige ganze rationale Funktion n-ten Grades mit reellen Koeffizienten. Ist dann $f(x)$ eine gegebene, im Intervalle

$-1 \leqq x \leqq +1$ stetige Funktion von x, dann heißt

$$(1) \qquad \int_{-1}^{+1} (f(x) - g(x))^2 \, dx$$

die im Sinne der ***Methode der kleinsten Quadrate*** genommene Abweichung von $g(x)$ von der Funktion $f(x)$ im Intervalle $(-1, +1)$, oder auch, kurz, die Besselsche Abweichung von $g(x)$ und $f(x)$.

Es gibt nun bekanntlich ein einziges Polynom $g(x)$, für welches die Abweichung (1) minimal wird, und dieses $g(x)$ erhalte ich, wenn ich $f(x)$ in die Legendresche Reihe

$$(2) \qquad f(x) \sim c_0 P_0(x) + c_1 P_1(x) + \ldots + c_n P_n(x) + \ldots$$

entwickle und von dieser unendlichen Entwicklung die Partialsumme mit dem Index n nehme; d. h.

$$(3) \qquad g(x) = c_0 P_0(x) + c_1 P_1(x) + \ldots + c_n P_n(x).$$

Hier bezeichnet $P_n(x)$ das n-te Legendresche Polynom.

Die durch unsere Minimumforderung gewonnene Folge von Polynomen

$$(4) \qquad g_0(x), g_1(x), \ldots, g_n(x), \ldots$$

konvergiert aber im allgemeinen nicht zu $f(x)$. Z. B. ist für die Funktion

$$(5) \qquad f(x) = \sum_{\nu=1}^{\infty} (P_{\nu^3}(x) - P_{\nu^3+2}(x)),$$

welche durch die an der rechten Seite stehenden, für $-1 \leqq x \leqq +1$ gleichmäßig und absolut konvergierenden Reihe definiert ist, die Folge (4) an der Stelle $x = +1$ augenscheinlich divergent.

Bilde ich aber, *wieder im Sinne der Methode der kleinsten Quadrate,*

$$(6) \qquad g_n^*(x) = \frac{g_0(x) + g_1(x) + \ldots + g_n(x)}{n+1},$$

und dann weiter

$$(7) \qquad g_n^{**}(x) = \frac{g_0^*(x) + g_1^*(x) + \ldots + g_n^*(x)}{n+1},$$

dann erhalte ich in der Folge der zweiten arithmetischen Mittel der Partialsummen der Legendreschen Reihe (2) von $f(x)$, d. h. in der Folge

$$(8) \qquad g_0^{**}(x), g_1^{**}(x), \ldots, g_n^{**}(x), \ldots,$$

die gewünschte, d. h. im Intervalle $-1 \leqq x \leqq +1$ gleichmäßig zu $f(x)$ konvergierende Folge von rationalen ganzen Funktionen. Es ist sogar noch

$$(9) \qquad m \leqq g_n^{**}(x) \leqq M,$$
$$-1 \leqq x \leqq +1, \qquad n = 0, 1, 2, \ldots, \infty,$$

wo m das Minimum, M das Maximum von $f(x)$ im Intervalle $-1 \leqq x \leqq +1$ bezeichnet.

Dieses Resultat habe ich im Jahre 1908 veröffentlicht[2]). Ich bemerke, daß bekanntlich T. H. Gronwall später bewiesen hat, daß sogar schon die Folge der ersten arithmetischen Mittel $g_n^*(x)$ unter (6) gleichmäßig zu $f(x)$ konvergiert, wenn auch für sie die Eigenschaft (9) verloren geht, die aber allerdings hier gar nicht gefordert ist. Weiter bemerke ich, daß, nach einer mündlichen Mitteilung von Professor Hilbert im Jahre 1902, Weierstraß selbst einen Beweis seines Approximationssatzes postulierte, der auf die Legendresche Entwicklung (2) der gegebenen Funktion $f(x)$ gegründet wäre.

§ 2.

Approximation einer stetigen Funktion durch eine rationale ganze Funktion mit Hilfe der Methode der kleinsten Quadrate, wenn $\frac{1}{\sqrt{1-x^2}}$ als Dichtigkeitsfunktion genommen wird, d. h. mit Benutzung der Teilsummen der Fourier-Tschebyscheffschen Reihe der zu approximierenden Funktion.

4. Wenn ich statt der Besselschen Abweichung (1) die Mehlersche

$$(10)\qquad \int_{-1}^{+1} \frac{1}{\sqrt{1-x^2}} (f(x) - g(x))^2 \, dx$$

minimisiere, so erhalte ich folgendes Resultat. Man bilde die Fouriersche Kosinusreihe von $f(\cos\theta)$

$$(11)\qquad f(\cos\theta) \sim c_0 + c_1 \cos\theta + \ldots + c_n \cos n\theta + \ldots,$$

d. h., in x,

$$(12)\qquad f(x) \sim c_0 T_0(x) + c_1 T_1(x) + \ldots + c_n T_n(x) + \ldots,$$

wo also $T_n(x)$ diejenige g. r. Funktion n-ten Grades bezeichnet, die $\cos n\theta$ durch $\cos\theta$ ausdrückt (das s. g. n-te Tschebyscheffsche Polynom), dann konvergieren im allgemeinen zwar nicht die Partialsummen

$$(13)\qquad g_n(x) = c_0 T_0(x) + c_1 T_1(x) + \ldots + c_n T_n(x),$$

wohl aber ihre arithmetischen Mittel

$$(14)\qquad g_n^*(x) = \frac{g_0(x) + g_1(x) + \ldots + g_n(x)}{n+1} \qquad (n = 0, 1, 2, \ldots, \infty)$$

gleichmäßig zu $f(x)$ im Intervalle $-1 \leqq x \leqq +1$, und es ist wieder $m \leqq g_n^*(x) \leqq M$.

Diese Methode habe ich im Jahre 1900 veröffentlicht[3]).

---

[2]) L. Fejér, Sur le développement d'une fonction arbitraire suivant les fonctions de Laplace, Comptes-Rendus, 3 février 1908.

[3]) L. Fejér, Sur les fonctions bornées et intégrables, Comptes Rendus, 10 décembre 1900.

Der Beweis des eben ausgesprochenen Approximationssatzes ist sehr einfach; der des § 1 ist etwas komplizierter, weil ich die merkwürdige Ungleichung

$$P_0(x) + P_1(x) + \dots + P_n(x) \geqq 0, \quad -1 \leqq x \leqq +1, \quad n = 0, 1, 2, \dots, \infty, \tag{15}$$

nicht entbehren kann. Aber ich wiederhole: es kommt mir nicht so sehr auf die Einfachheit des Beweises, als auf die des *Herstellungsprinzips* und der *Herstellungsregel* an, welche zur Bestimmung der Näherungspolynome dienen.

5. Man kann beanstanden, daß sowohl im definierenden Prinzipe, als in der Herstellungsregel des § 1 und § 2 die *Integration* eine Rolle spielt. In der nun folgenden Interpolationsmethode tritt die Integration nicht auf. Ich betrachte hier nicht jenes Polynom n-ten Grades $g(x)$, für welches die Integralabweichung von $f(x)$ in diesem oder jenem Sinne möglichst klein ist, sondern ein Interpolationspolynom, welches also an $(n+1)$ Stellen x mit $f(x)$ geradezu gleich ist. Will man aber gleichmäßige Konvergenz im Intervalle $-1 \leqq x \leqq +1$, so kann man, nach einem wichtigen Satze von G. Faber[4]), bekanntlich mit den Lagrangeschen Parabeln nicht auskommen. Ich betrachte also höhere Interpolationsgebilde, s. g. Hermitesche Parabeln für geeignet gewählte Abszissengruppen; diese bilden ein Analogon zu den arithmetischen Mitteln der früher erwähnten Orthogonalreihen und konvergieren gleichmäßig zur stetigen Funktion $f(x)$.

§ 3.

Die Lagrangesche Interpolationsformel.

6. Es seien $x_1, x_2, \dots, x_n$ voneinander verschiedene Abszissen und $y_1, y_2, \dots, y_n$ Ordinatenwerte; dann lautet die Lagrangesche Formel

$$L(x) = \sum_{k=1}^{n} y_k l_k(x), \tag{16}$$

wo

$$l_k(x) = \frac{\omega(x)}{\omega'(x_k)(x - x_k)}, \tag{17}$$

$$\omega(x) = C(x - x_1)(x - x_2)\dots(x - x_n) \tag{18}$$

und $C \neq 0$ eine willkürliche Konstante bezeichnet.

Sie stellt diejenige ganze rationale Funktion von höchstens $(n-1)$-tem Grade dar, die an den Stellen $x_1, x_2, \dots, x_n$ der Reihe nach die Werte

[4]) G. Faber, Über die interpolatorische Darstellung stetiger Funktionen, Jahresbericht der deutschen Mathematikervereinigung (1914), S. 192—210.

$y_1, y_2, \ldots, y_n$ annimmt. (Es gibt immer eine, aber auch nur eine solche g. r. Funktion.) Dies alles läßt sich leicht beweisen.

Die Polynome $l_k(x)$ nenne ich die *Grundpolynome der Lagrangeschen Interpolation*. Setzt man in die Lagrangesche Formel $y_1 = y_2 = \ldots = y_n = 1$, so erhält man auf Grund des Vorhergehenden

$$\sum_{k=1}^{n} l_k(x) = 1. \tag{19}$$

§ 4.

Die Hermitesche Interpolationsformel.

7. Es seien $x_1, x_2, \ldots, x_n$ wieder n voneinander verschiedene Abszissenwerte und

$$y_1, y_2, \ldots, y_n, \qquad y'_1, y'_2, \ldots, y'_n \tag{20}$$

weitere $2n$ gegebene Werte; dann lautet die Hermitesche Interpolationsformel

$$\mathsf{X}(x) = \sum_{k=1}^{n} y_k h_k(x) + \sum_{k=1}^{n} y'_k \mathfrak{h}_k(x), \tag{21}$$

wo

$$h_k(x) = \left(1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x - x_k)\right)(l_k(x))^2 = v_k(x)(l_k(x))^2, \tag{22}$$

$$\mathfrak{h}_k(x) = (x - x_k)(l_k(x))^2 \tag{23}$$

bezeichnet. Es ist hier wieder

$$\omega(x) = C(x - x_1)(x - x_2) \ldots (x - x_n) \qquad (C \neq 0), \tag{24}$$

$$l_k(x) = \frac{\omega(x)}{\omega'(x_k)(x - x_k)}, \tag{25}$$

und $v_k(x)$ ist augenscheinlich die abgekürzte Bezeichnung des „charakteristischen Linearfaktors" in $h_k(x)$, d. h.

$$v_k(x) = 1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x - x_k). \tag{26}$$

Die Hermitesche Interpolationsformel (21) stellt diejenige ganze rationale Funktion von höchstens $(2n-1)$-tem Grade dar, die an den Stellen $x_1, x_2, \ldots, x_n$ der Reihe nach die Werte $y_1, y_2, \ldots, y_n$ und deren Ableitung an denselben Stellen der Reihe nach die Werte $y'_1, y'_2, \ldots, y'_n$ annimmt. Es gibt immer ein solches Polynom, aber auch nur ein einziges. Dies läßt sich alles leicht beweisen.

Die n Polynome unter (22)

$$h_1(x), h_2(x), \ldots, h_n(x) \tag{27}$$

nenne ich die *Grundpolynome erster Art* und die n Polynome unter (23)

(28) $$\mathfrak{h}_1(x), \mathfrak{h}_2(x), \ldots, \mathfrak{h}_n(x)$$

nenne ich die *Grundpolynome zweiter Art, der Hermiteschen Interpolation* (21).

8. Setzt man in die Hermitesche Formel (21)

$$y_1 = y_2 = \ldots = y_n = 1, \quad y_1' = y_2' = \ldots = y_n' = 0,$$

so erhält man für die Grundpolynome erster Art die identische Relation

(29) $$\sum_{k=1}^{n} h_k(x) \equiv 1.$$

Setzt man in (21)

$$y_1 = y_2 = \ldots = y_n = 0, \quad y_1' = y_2' = \ldots = y_n' = 1,$$

so erhält man

(30) $$\sum_{k=1}^{n} \mathfrak{h}_k(x) \equiv \mathfrak{h}(x),$$

wo $\mathfrak{h}(x)$ eben dasjenige Polynom $(2n-1)$-ten Grades bezeichnet, welches an den Stellen $x_1, x_2, \ldots, x_n$ verschwindet und an allen diesen Stellen die Ableitung 1 hat.

§ 5.

Die Treppenparabel und die Wellenparabel.

9. Die Hermitesche Formel (21) zerlegt das zu den Daten $y_1, y_2, \ldots, y_n$, $y_1', y_2', \ldots, y_n'$ gehörige Hermitesche Interpolationspolynom in zwei Summanden:

(31) $$X(x) = H(x) + \mathfrak{H}(x),$$

wo

(32) $$H(x) = \sum_{k=1}^{n} y_k h_k(x),$$

(33) $$\mathfrak{H}(x) = \sum_{k=1}^{n} y_k' \mathfrak{h}_k(x).$$

Die Polynome $H(x)$ und $\mathfrak{H}(x)$ lassen sich aber auch unabhängig von ihren formelmäßigen Darstellungen (32) und (33) in folgender Weise charakterisieren:

1. $H(x)$ ist dasjenige Polynom von höchstens $(2n-1)$-tem Grade, welches an den Stellen $x_1, x_2, \ldots, x_n$ die Werte $y_1, y_2, \ldots, y_n$ annimmt, und dessen Derivierte an diesen Stellen verschwindet. Ich nenne dieses Polynom: das *Treppenpolynom* [die Kurve $y = H(x)$ nenne ich die durch die Punkte $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ gehende *Treppenparabel*].

2. $\mathfrak{H}(x)$ ist dasjenige Polynom von höchstens $(2n-1)$-tem Grade, welches an den Stellen $x_1, x_2, \ldots, x_n$ verschwindet und dessen Derivierte an diesen Stellen die Werte $y_1', y_2', \ldots, y_n'$ besitzt. Ich kann dieses Polynom vielleicht das *Wellenpolynom* nennen [und die Parabel $y = \mathfrak{H}(x)$ die durch die Punkte $(x_1, 0), (x_2, 0), \ldots, (x_n, 0)$ gehende und die „Neigungen" $y_1', y_2', \ldots, y_n'$ besitzende *Wellenparabel*].

§ 6.

Das Vorzeichen der Grundpolynome bei der Hermiteschen Interpolation. Die konjugierten Punkte.

10. Von nun an will ich die voneinander immer verschiedenen Abszissen $x_1, x_2, \ldots, x_n$ ausdrücklich reell annehmen und lenke meine Aufmerksamkeit auf das *Vorzeichen der Grundpolynome.*

Ein Blick auf die Formel (17) des Lagrangeschen Grundpolynoms $l_k(x)$ lehrt, daß dieses sein Vorzeichen $(n-1)$-mal ändert, wenn x von $-\infty$ bis $+\infty$ variiert.

Einen ganz anderen Sachverhalt finden wir aber bei den Hermiteschen Grundpolynomen vor.

1. Das Grundpolynom *erster Art* $h_k(x)$ unter (22) hat *höchstens eine Zeichenwechselstelle* und zwar dort, wo

$$v_k(x) = 1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x - x_k) = 0 \tag{34}$$

ist, d. h. an der Stelle

$$X_k = x_k + \frac{\omega'(x_k)}{\omega''(x_k)}. \tag{35}$$

Ich nenne diese Zeichenwechselstellen $X_1, X_2, \ldots, X_n$ die zu den Stellen $x_1, x_2, \ldots, x_n$ gehörigen *konjugierten Stellen.* (Ist für ein k der Wert $\omega''(x_k) = 0$, so ist $X_k = \infty$; in diesem Falle ist $v_k(x) \equiv 1$ und $h_k(x)$ hat überhaupt keine Zeichenwechselstelle.) Die Lage der konjugierten Stellen spielt im folgenden eine entscheidende Rolle. Es ist übrigens

$$X_k = x_k + \frac{1}{2\left(\frac{1}{x_k - x_1} + \ldots + \frac{1}{x_k - x_{k-1}} + \frac{1}{x_k - x_{k+1}} + \ldots + \frac{1}{x_k - x_n}\right)} \tag{36}$$
$$(k = 1, 2, \ldots, n)$$

eine Formel, die wir aber im folgenden nicht gebrauchen werden.

2. Das Grundpolynom *zweiter Art* $\mathfrak{h}_k(x)$ unter (23) hat *eine* Zeichenwechselstelle, nämlich die Stelle

$$x = x_k. \tag{37}$$

§ 7.
Ein Satz über die Treppenparabel.

11. Die Stellen $x_1, x_2, \ldots, x_n$ sollen nun in das Intervall $-1 \leqq x \leqq +1$ fallen. Ich nehme weiter an, daß die konjugierten Punkte $X_1, X_2, \ldots, X_n$ *außerhalb* dieses Intervalles liegen (d. h. genauer, entweder sei $X_k \geqq 1$ oder $X_k \leqq -1$). Dann sind, mit Rücksicht auf $v_k(x_k) = 1$, die Grundpolynome erster Art $h_k(x)$ unter (22) alle *nichtnegativ* für $-1 \leqq x \leqq +1$. Daraus ergibt sich mit Rücksicht auf (29) und (32) der folgende

Satz. Sind für die im Intervall $-1 \leqq x \leqq +1$ liegenden Punkte $x_1, x_2, \ldots, x_n$ die konjugierten Punkte $X_1, X_2, \ldots, X_n$ nicht im Innern dieses Intervalls gelegen, dann liegen die Werte des Treppenpolynoms $H(x) = \sum_{k=1}^{n} y_k h_k(x)$ für $-1 \leqq x \leqq +1$ zwischen dem kleinsten und dem größten der Werte $y_1, y_2, \ldots, y_n$.

§ 8.
Neue Charakterisierung der Legendre-Gaußschen und Tschebyscheffschen Abszissen, und zwar mit Hilfe der konjugierten Punkte.

12. Im folgenden will ich nun die Abszissen $x_1, x_2, \ldots, x_n$ ganz speziell wählen.

A. Der konjugierte Punkt X_k liege im Mittelpunkt der Strecke, die von x_k bis zum konjugiert *harmonischen* Punkte von x_k reicht, wo -1 und $+1$ als Grundpunkte gedacht sind, d. h.

$$X_k = \frac{x_k + \frac{1}{x_k}}{2} \qquad (k = 1, 2, \ldots, n). \tag{38}$$

B. Der konjugierte Punkt X_k liege im konjugierten harmonischen Punkte von x_k, wo wieder -1 und $+1$ als Grundpunkte gedacht sind, d. h.

$$X_k = \frac{1}{x_k} \qquad (k = 1, 2, \ldots, n). \tag{39}$$

In beiden Fällen liegt X_k nicht im Innern des Intervalles $(-1, +1)$, so daß also nach Nr. 11 sämtliche Grundpolynome erster Art $h_k(x)$ im Intervall $-1 \leqq x \leqq +1$ nichtnegativ ausfallen.

Bestimmen wir nun in beiden Fällen die unseren Forderungen entsprechenden Abszissen $x_1, x_2, \ldots, x_n$, oder, was dasselbe ist, die ganze rationale Funktion n-ten Grades $\omega(x)$.

13. Laut (35) fordern wir im Falle A, daß

$$\frac{x_k + \frac{1}{x_k}}{2} = x_k + \frac{\omega'(x_k)}{\omega''(x_k)} \tag{40}$$

erfüllt sei, für $k = 1, 2, \ldots, n$, d. h. daß

$$(41) \qquad (1 - x_k^2)\,\omega''(x_k) - 2\,x_k\,\omega'(x_k) = 0$$

sei für $k = 1, 2, \ldots, n$. Wir fordern also, daß das Polynom von höchstens n-tem Grade

$$(1 - x^2)\,\omega''(x) - 2\,x\,\omega'(x)$$

an sämtlichen Nullstellen $x_1, x_2, \ldots, x_n$ des Polynoms $\omega(x)$ verschwinde. Es muß also

$$(42) \qquad (1 - x^2)\,\omega''(x) - 2\,x\,\omega'(x) = \text{Konst.}\,\omega(x)$$

sein, woraus wir durch Vergleichung der Koeffizienten von x^n auf beiden Seiten der Gleichung (42) sofort ersehen, daß die Konstante auf der rechten Seite von (42) gleich $-n(n+1)$ ist. Also gilt schließlich

$$(43) \qquad (1 - x^2)\,\omega''(x) - 2\,x\,\omega'(x) + n(n+1)\,\omega(x) = 0\,.$$

Das ist die Differentialgleichung des n-ten Legendreschen Polynoms. Also sind im Falle A die $x_1, x_2, \ldots, x_n$ die s. g. Legendreschen oder Gaußschen Abszissen, d. h. die Wurzeln der Gleichung $P_n(x) = 0$, wo $P_n(x)$ das n-te Legendresche Polynom bezeichnet. Wir haben also das folgende Resultat erhalten:

Für die Wurzeln der Gleichung $P_n(x) = 0$, und nur für diese Punktgruppe, liegt der konjugierte Punkt X_k von x_k für jeden der Werte $k = 1, 2, \ldots, n$ genau in der Mitte der Strecke von x_k bis zum konjugiert harmonischen Punkte von x_k, wenn -1 und $+1$ als Grundpunkte bebetrachtet werden.

14. Durch die Forderung B

$$(44) \qquad \frac{1}{x_k} = x_k + \frac{\omega'(x_k)}{\omega''(x_k)}$$

erhalten wir für $\omega(x)$ auf demselben Wege die Differentialgleichung

$$(45) \qquad (1 - x^2)\,\omega''(x) - x\,\omega'(x) + n^2\,\omega(x) = 0\,,$$

welchem das n-te Tschebyscheffsche Polynom genügt.

Für die Wurzeln der Gleichung $T_n(x) = \cos(n \arccos x) = 0$, d. h. für die Punktgruppe

$$(46) \qquad x_k = \cos(2\,k - 1)\frac{\pi}{2\,n} \qquad (k = 1, 2, \ldots, n),$$

und nur für diese Punktgruppe liegt der konjugierte Punkt X_k von x_k für jeden der Werte $k = 1, 2, \ldots, n$ im konjugiert harmonischen Punkte von x_k, wo wieder -1 und $+1$ als Grundpunkte für das harmonische Paar gedacht sind.

Die Tschebyscheffschen Abszissen lassen sich übrigens so konstruieren: man schlage über der Strecke von -1 bis $+1$ als Durchmesser einen Halbkreis, teile diesen in n gleiche Teilbögen und projiziere die n Mittelpunkte dieser Teilbögen auf den Durchmesser von -1 bis $+1$. Die Projektionspunkte sind die Tschebyscheffschen Abszissenpunkte.

§ 9.

Charakterisierung der allgemeinen ultrasphärischen Abszissen mit Hilfe der konjugierten Punkte.

15. Ich stelle drittens die Forderung:

C. Die konjugierte Stelle X_k sei

$$X_k = \frac{\lambda x_k + \frac{1}{x_k}}{\lambda + 1} \qquad (0 \leqq \lambda \leqq 1, \quad k = 1, 2, \ldots, n),$$

d. h. der konjugierte Punkt teile die Strecke von x_k bis zu seinem konjugiert harmonischen Paare $\frac{1}{x_k}$ in zwei Teile, deren *Verhältnis*, für jeden der Werte $k = 1, 2, \ldots, n$, dieselbe Größe λ habe, $(0 \leqq \lambda \leqq 1)$. D. h.

$$\frac{\frac{1}{x_k} - X_k}{X_k - x_k} = \lambda \qquad (0 \leqq \lambda \leqq 1, \quad k = 1, 2, \ldots, n).$$

Diese Forderung führt, auf dem angegebenen Wege, sofort zur Differentialgleichung der ultrasphärischen Polynome[8])

$$(1 - x^2)\,\omega'' - (\lambda + 1)\,x\,\omega' + n(n + \lambda)\,\omega = 0.$$

Der Legendresche Fall A entspricht dem Werte $\lambda = 1$, der Tschebyscheffsche Fall B dem Werte $\lambda = 0$.

Ich bemerke, daß die zum Verhältniswerte λ $(\lambda > 0)$ gehörigen ultrasphärischen Polynome $\omega_n(x)$ in der Potenzreihenentwicklung nach z der erzeugenden Funktion

$$\frac{1}{(1 - 2xz + z^2)^{\frac{\lambda}{2}}} = \omega_0(x) + \omega_1(x)\,z + \ldots + \omega_n(x)\,z^n + \ldots$$

als Koeffizienten auftreten, und daß hier die charakteristische Linearfunktion unter (26) die Form

$$v_k(x) = \frac{1 - (\lambda + 1)\,x\,x_k + \lambda x_k^2}{1 - x_k^2}$$

hat.

---

[8]) N. Nielsen, Théorie des fonctions métasphériques, Paris, Gauthier-Villars, 1911. E. Kogbetliantz, Recherches sur les séries ultrasphériques, Journal de Mathématiques pures et appliquées (9) **3** (1924), p. 107—187.

Endlich erwähne ich hier, daß im Falle der Newtonschen Abszissen

$$x_k = -1 + k\frac{2}{n+1} \qquad (k = 1, 2, \ldots, n)$$

sämtliche konjugierte Punkte X_k in das Intervall $(-1, +1)$ fallen. Eine Ausnahme kommt nur vor, wenn n ungerade ist; dann fällt der der sodann auftretenden Interpolationsstelle $x = 0$ entsprechende konjugierte Punkt ins Unendliche.

16. Die Nummern 12 bis 15 enthalten eine neue Kennzeichnung der ultrasphärischen Abszissen. Es führte zu ihnen die Betrachtung der konjugierten Punkte der Hermiteschen Interpolation. In den §§ 1, 2 sind diese speziellen Abszissengruppen auf einem ganz anderen Wege aufgetreten, nämlich durch die klassische Methode der kleinsten Quadrate. Schließlich erinnere ich noch an einen dritten Weg, auf dem die ultrasphärischen Abszissen unmittelbar hervortreten: es ist dies die klassische Frage nach gewissen günstigsten mechanischen Quadraturen (Gauß, Mehler). Im Reste dieser Arbeit beschränke ich mich nun ausschließlich auf den Grenzfall $\lambda = 0$, d. h. auf die Tschebyscheffschen Abszissen, ohne meine Resultate, die sich auf allgemeine ultrasphärische Abszissen beziehen, zu erwähnen.

§ 10.

Über die wichtigsten Eigenschaften der Grundfunktionen beider Art im Falle Tschebyscheffscher Abszissen.

17. Es seien mit $x_1, x_2, \ldots, x_n$ die Tschebyscheffschen Abszissen bezeichnet.

Ich erinnere zunächst daran, daß immer

$$\text{(I)} \qquad \sum_{k=1}^{n} h_k(x) = 1$$

ist.

Eine leichte Rechnung zeigt, daß im Falle Tschebyscheffscher Abszissen

$$\text{(II)} \qquad \sum_{k=1}^{n} \mathfrak{h}_k(x) = \mathfrak{h}(x) = \mathfrak{h}(\cos\theta) = \frac{1}{n}\cos n\theta \cos(n-1)\theta = \frac{1}{2n}(\cos\theta + \cos(2n-1)\theta)$$

ist. Diese interessante Gleichung habe ich hier angeführt, werde sie aber beim Konvergenzbeweis nicht benützen.

Wegen $h_k(x) \geqq 0$, für $-1 \leqq x \leqq +1$, ist natürlich auch

$$\text{(III)} \qquad \sum_{k=1}^{n} |h_k(x)| = 1.$$

Um die anderen, zum Konvergenzbeweise nötigen Abschätzungen erhalten zu können, muß ich die Grundpolynome im vorliegenden Tschebyscheffschen Falle etwas näher betrachten. Laut (22) und (23) ist immer

$$h_k(x) = \left(1 - \frac{\omega''(x_k)}{\omega'(x_k)}(x - x_k)\right)(l_k(x))^2 = v_k(x)(l_k(x))^2, \tag{47}$$

$$\mathfrak{h}_k(x) = (x - x_k)(l_k(x))^2. \tag{48}$$

Im Falle Tschebyscheffscher Abszissen ist aber, auf Grund der Gleichung (44),

$$\frac{\omega''(x_k)}{\omega'(x_k)} = \frac{x_k}{1 - x_k^2}, \tag{49}$$

und, wegen

$$\omega(x) = T(x) = T(\cos\theta) = \cos n\theta, \tag{50}$$

$$(\omega'(x_k))^2 = \frac{n^2}{1 - x_k^2}. \tag{51}$$

Also ist, mit Rücksicht auf (47) bis (51),

$$h_k(x) = v_k(x)(l_k(x))^2 = \frac{1 - x x_k}{1 - x_k^2}(l_k(x))^2 = \frac{1}{n^2}(1 - x x_k)\left(\frac{T(x)}{x - x_k}\right)^2 \tag{52}$$

und

$$\mathfrak{h}_k(x) = (x - x_k)(l_k(x))^2 = \frac{x - x_k}{v_k(x)} h_k(x) = (x - x_k)\cdot\frac{1 - x_k^2}{1 - x x_k} h_k(x), \tag{53}$$

oder mit anderer Gruppierung der Faktoren

$$\mathfrak{h}_k(x) = \frac{x - x_k}{1 - x x_k}(1 - x_k^2) h_k(x). \tag{53'}$$

Aus (52) erhalten wir zunächst, mit Rücksicht auf $|T(x)| = |T(\cos\theta)| = |\cos n\theta| \leqq 1$,

$$|h_k(x)| = h_k(x) \leqq \frac{2}{(x - x_k)^2}\cdot\frac{1}{n^2}, \tag{54}$$

für $-1 \leqq x \leqq +1,\ x \neq x_k$.

Da weiter

$$\left|\frac{x - x_k}{1 - x x_k}\right| \leqq 1, \quad \text{für} \quad |x| \leqq 1, \tag{55}$$

so ist, mit Rücksicht auf $|1 - x_k^2| = 1 - x_k^2 \leqq 1$ und auf Grund der Formel (53′),

$$|\mathfrak{h}_k(x)| \leqq h_k(x), \quad \text{für} \quad |x| \leqq 1, \tag{56}$$

eine, im Tschebyscheffschen Falle gültige, wichtige Ungleichung.

Durch Summation erhalte ich nun aus (56), mit Rücksicht auf (I),

$$\sum_{k=1}^{n} |\mathfrak{h}_k(x)| \leqq \sum_{k=1}^{n} h_k(x) = 1. \tag{IV}$$

18. Die feste Stelle x sei durch ein festes Intervall $(x-\varepsilon,\ x+\varepsilon)$ umgeben. Diesem Intervalle entsprechend teilen wir die Summen $\sum\limits_{k=1}^{n}|h_k(x)|$ und $\sum\limits_{k=1}^{n}|\mathfrak{h}_k(x)|$ in zwei Teile:

$$\sum_{k=1}^{n}|h_k(x)| = \sum_{k=1}^{n} h_k(x) = \sum{}' h_k(x) + \sum{}'' h_k(x), \tag{57}$$

$$\sum_{k=1}^{n}|\mathfrak{h}_k(x)| = \sum{}' |\mathfrak{h}_k(x)| + \sum{}'' |\mathfrak{h}_k(x)|, \tag{58}$$

wo $\sum'$ immer eine Summe bedeutet, die über diejenigen Indizes k zu erstrecken ist, denen ein in das abgeschlossene Intervall $(x-\varepsilon,\ x+\varepsilon)$ fallendes x_k entspricht. $\sum''$ bezeichnet immer eine Summe, die über allen übrigen Indizes k zu erstrecken ist. (Die eine oder andere Summe kann unter Umständen auch leer ausfallen.)

Nun ist, mit Rücksicht auf (I) und $h_k(x) \geqq 0$,

$$\sum{}' |h_k(x)| = \sum{}' h_k(x) \leqq \sum_{k=1}^{n} h_k(x) = 1, \tag{V}$$

und, mit Rücksicht auf (IV),

$$\sum{}' |\mathfrak{h}_k(x)| \leqq \sum{}' h_k(x) \leqq 1. \tag{VI}$$

Weiter ist, mit Rücksicht auf (54),

$$\sum{}'' |h_k(x)| = \sum{}'' h_k(x) \leqq \frac{2}{\varepsilon^2} \sum{}'' \frac{1}{n^2} \leqq \frac{2}{\varepsilon^2} \sum_{k=1}^{n} \frac{1}{n^2} = \frac{2}{\varepsilon^2} n \cdot \frac{1}{n^2} = \frac{2}{\varepsilon^2} \frac{1}{n}, \tag{VII}$$

und, mit Rücksicht auf (56) und (VII),

$$\sum{}'' |\mathfrak{h}_k(x)| \leqq \sum{}'' h_k(x) \leqq \frac{2}{\varepsilon^2} \frac{1}{n}. \tag{VIII}$$

Ich kann aber außerdem noch zeigen, daß im Falle Tschebyscheffscher Abszissen

$$\lim_{n=\infty} \sum_{k=1}^{n} |\mathfrak{h}_k(x)| = 0 \tag{IX}$$

gültig ist, und zwar gleichmäßig im Intervalle $-1 \leqq x \leqq +1$. Da nämlich

$$\frac{1-x_k^2}{1-x x_k} \leqq \frac{1-|x_k|^2}{1-|x_k|} = 1 + |x_k| \leqq 2,$$

so gilt, mit Rücksicht auf (53), für $|\mathfrak{h}_k(x)|$ außer der Ungleichung (56) auch noch die ebenfalls wichtige Ungleichung

$$|\mathfrak{h}_k(x)| \leqq 2\,|x - x_k|\, h_k(x), \tag{X}$$

aus welcher, mit Rücksicht auf (V), die Abschätzung

$$\sum{}' |\mathfrak{h}_k(x)| \leqq 2 \sum{}' |x - x_k|\, h_k(x) \leqq 2\varepsilon \sum{}' h_k(x) \leqq 2\varepsilon \tag{VI$'$}$$

geschlossen werden kann.

Aus (VI′), (VIII) und (58) folgt nun

$$\sum_{k=1}^{n} |\mathfrak{h}_k(x)| \leqq 2\varepsilon + \frac{2}{\varepsilon^2}\,\frac{1}{n} \leqq 3\varepsilon,$$

wenn nur n gehörig groß ist. Da ε eine beliebig kleine positive Größe bezeichnet, so habe ich (IX) schon bewiesen, d. h. das folgende Resultat erhalten:

Die Summe der absoluten Beträge der Grundpolynome zweiter Art $\mathfrak{h}_k(x)$ konvergiert im Falle Tschebyscheffscher Abszissen zu Null, wenn $n \to \infty$, und zwar gleichmäßig im Intervalle $-1 \leqq x \leqq +1$.

Beim Konvergenzbeweis werde ich nicht (IV), sondern (IX) gebrauchen.

§ 11.

Beweis der gleichmäßigen Konvergenz der Hermiteschen Interpolationspolynome zur Funktion, wenn ihre Steilheit in den Tschebyscheffschen Abszissen beschränkt bleibt.

19. Es sei nun $y = f(x)$ eine im Intervalle $-1 \leqq x \leqq +1$ überall stetige Funktion. Es seien $x_1, x_2, \ldots, x_n$ die zum Index n gehörigen Tschebyscheffschen Abszissen. Es seien weiter $y_1, y_2, \ldots, y_n$ die Werte von $y = f(x)$ an den Stellen $x_1, x_2, \ldots, x_n$, und $y_1', y_2', \ldots, y_n'$ seien *beliebige n Werte, die dem absoluten Betrage nach nicht größer sind als Δ, wo Δ eine beliebig gewählte nichtnegative absolute Konstante bezeichnet, d. h. es sei*

$$y_k = f(x_k), \quad |y_k'| \leqq \Delta \qquad (k = 1, 2, \ldots, n). \tag{59}$$

(Den oberen Index n, den ich bei allen Größen

$$\begin{matrix} x_1, & x_2, & \ldots, & x_n \\ y_1, & y_2, & \ldots, & y_n \\ y_1', & y_2', & \ldots, & y_n' \end{matrix} \tag{60}$$

ansetzen müßte, kann ich ruhig weglassen, ohne ein Mißverständnis zu befürchten.)

Das zu den so definierten Daten (60) gehörige Hermitesche Interpolationspolynom lautet

$$\mathrm{X}_n(x) = \sum_{k=1}^{n} y_k h_k(x) + \sum_{k=1}^{n} y_k' \mathfrak{h}_k(x) = H_n(x) + \mathfrak{H}_n(x). \tag{61}$$

Ich behaupte nun, daß das *Wellenpolynom*

$$\mathfrak{H}_n(x) = \sum_{k=1}^{n} y_k' \mathfrak{h}_k(x) \tag{62}$$

mit $n \to \infty$ für $-1 \leqq x \leqq +1$ *gleichmäßig zu* 0 *konvergiert*. Tatsächlich ist

$$(63)\qquad |\mathfrak{H}_n(x)| \leqq \sum_{k=1}^{n} |y_k'| \, |\mathfrak{h}_k(x)| \leqq \Delta \sum_{k=1}^{n} |\mathfrak{h}_k(x)| \to 0$$

auf Grund der Limesgleichung (IX).

Es bleibt also nur noch übrig das Treppenpolynom

$$(64)\qquad H_n(x) = \sum_{k=1}^{n} y_k h_k(x)$$

für $n \to \infty$ zu untersuchen. Es sei x eine beliebige Stelle des Intervalls $-1 \leqq x \leqq +1$, und y bezeichne den Wert von $f(x)$ an der Stelle x. Dann ist, mit Rücksicht auf (64) und (I),

$$(65)\qquad H_n(x) - y = \sum_{k=1}^{n} (y_k - y) h_k(x).$$

Es sei nun δ eine beliebige positive Größe, und ε so bestimmt, daß

$$(66)\qquad |f(t) - f(x)| \leqq \delta,$$

wenn

$$(67)\qquad |t - x| \leqq \varepsilon,$$

wo auch der Wert x im Intervalle $-1 \leqq x \leqq +1$ gelegen sei. Aus (65), (66) und (67) folgt, mit Berücksichtigung von (V) und (VII), unmittelbar

$$(68)\qquad \begin{aligned} |H_n(x) - y| &\leqq \sum_{k=1}^{n} |y_k - y| h_k(x) \\ &= \sum{}' |y_k - y| h_k(x) + \sum{}'' |y_k - y| h_k(x) \\ &\leqq \delta \sum{}' h_k(x) + S \sum{}'' h_k(x) \leqq \delta + S \cdot \frac{2}{\varepsilon^2} \cdot \frac{1}{n}, \end{aligned}$$

wo S die größte Schwankung von $f(t)$ im Intervalle $-1 \leqq t \leqq +1$ bezeichnet. Also ist

$$(69)\qquad |H_n(x) - y| \leqq 2\delta,$$

wenn nur n gehörig groß ist, und zwar im ganzen Intervalle $-1 \leqq x \leqq +1$. Hiermit ist aber bewiesen, daß im Intervalle $-1 \leqq x \leqq +1$ das Treppenpolynom $H_n(x)$ gleichmäßig zu $f(x)$ konvergiert, d. h.

$$(70)\qquad \lim_{n=\infty} H_n(x) = f(x)$$

stattfindet. Ich habe soeben bewiesen, daß für $-1 \leqq x \leqq +1$ gleichmäßig

$$(71)\qquad \lim_{n=\infty} \mathfrak{H}_n(x) = 0$$

stattfindet. Also ist, mit Rücksicht auf (61),

$$(72)\qquad \lim_{n=\infty} X_n(x) = f(x),$$

und zwar gleichmäßig im Intervalle $-1 \leqq x \leqq +1$.

Das so erhaltene Theorem läßt sich in folgender Weise formulieren.

Es sei $y = f(x)$ eine beliebig gegebene, im Intervalle $-1 \leqq x \leqq +1$ überall stetige Funktion. $x_1, x_2, \ldots, x_n$ sollen die Tschebyscheffschen Abszissen bezeichnen, d. h. es sei

$$x_k = \cos(2k-1)\frac{\pi}{2n} \qquad (k = 1, 2, \ldots, n),$$

und $y_1, y_2, \ldots, y_n$ sollen die Werte bezeichnen, die die Funktion $f(x)$ an diesen Stellen, der Reihe nach, annimmt, d. h. es sei

$$y_1 = f(x_1), \quad y_2 = f(x_2), \quad \ldots, \quad y_n = f(x_n).$$

Durch die Punkte

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n),$$

als Knotenpunkte, geht nun eine n-fach unendliche Schar von Parabeln von höchstens $(2n-1)$-tem Grade. Es sei $y = \mathrm{X}_n(x)$ irgendeine Parabel dieser Schar, für welche jedoch die Richtungstangenten in allen Knotenpunkten dem absoluten Betrage nach nicht größer sind als Δ, d. h.

$$|\mathrm{X}_n'(x_1)| \leqq \Delta, \quad |\mathrm{X}_n'(x_2)| \leqq \Delta, \quad \ldots, \quad |\mathrm{X}_n'(x_n)| \leqq \Delta,$$

wo Δ eine beliebig vorgeschriebene numerische Konstante bezeichnet). Dann ist*

$$\lim_{n=\infty} \mathrm{X}_n(x) = f(x),$$

und zwar gleichmäßig im Intervalle $-1 \leqq x \leqq +1$.

Für die Interpolationsparabeln $2n$-ten Grades ist dieser Satz *nicht* richtig.

§ 12.

Spezielle Fälle.

20. Nimmt man im vorigen allgemeinen Theoreme, für jedes n, speziell

$$y_1' = y_2' = \ldots = y_n' = 0,$$

so erhält man, daß die von mir Treppenparabeln genannten speziellen Hermiteschen Parabeln für $-1 \leqq x \leqq +1$ gleichmäßig zu $y = f(x)$ konvergieren, d. h. den folgenden Satz (der sich, wie aus dem Vorhergehenden ersichtlich, kürzer direkt beweisen läßt):

*) Über weitergehende Resultate möchte ich an anderer Stelle berichten. Hier bemerke ich nur, daß, allgemeiner, $\mathrm{X}_n(x)$ immer mit $\lim n = \infty$ im Intervalle $-1 \leqq x \leqq +1$ gleichmäßig zu $f(x)$ konvergiert, wenn nur der größte unter den n Quotienten

$$|\mathrm{X}_n'(x_k)|\sqrt{1-x_k^2} : \frac{n}{\log n} \qquad (k = 1, 2, \ldots, n),$$

mit $\lim n = \infty$ zu 0 konvergiert.

Bezeichnet $H_n(x)$ diejenige ganze rationale Funktion von höchstens $(2n-1)$-tem Grade in x, die an den Tschebyscheffschen Abszissen $x_1, x_2, \ldots, x_n$ denselben Wert annimmt wie die im Intervalle $-1 \leqq x \leqq +1$ beliebig gegebene überall stetige Funktion $f(x)$, deren Derivierte aber an allen diesen Stellen verschwindet, so ist $\lim_{n=\infty} H_n(x) = f(x)$, und zwar gleichmäßig im Intervalle $-1 \leqq x \leqq +1$. In Formeln:

$$\lim_{n=\infty} \frac{1}{n^2} \sum_{k=1}^{n} f(x_k^{(n)})(1 - x x_k^{(n)}) \left(\frac{T_n(x)}{x - x_k^{(n)}}\right)^2 = f(x),$$

und zwar gleichmäßig für $-1 \leqq x \leqq +1$, *wo*

$$T_n(x) = \cos n(\arccos x),$$

$$x_k^{(n)} = \cos(2k-1)\frac{\pi}{2n} \qquad (k = 1, 2, \ldots, n).$$

Während also das Lagrangesche Interpolationspolynom $L_n(x)$ — diejenige ganze rationale Funktion von höchstens $(n-1)$-tem Grade, die an den Tschebyscheffschen Stellen $x_1, x_2, \ldots, x_n$ mit unserer Funktion $f(x)$ übereinstimmt — nicht notwendigerweise gleichmäßig zu $f(x)$ konvergiert, konvergiert das Treppenpolynom $H_n(x)$ — das man auch als dasjenige Polynom von höchstens $(2n-1)$-tem Grade charakterisieren kann, welches an den besagten Stellen $x_1, x_2, \ldots, x_n$, der Reihe nach, die Werte

$$y_1 = f(x_1), \quad y_2 = f(x_2), \quad \ldots, \quad y_n = f(x_n)$$

doppelt annimmt — für $\lim n = \infty$ immer gleichmäßig zu $f(x)$, im Intervalle $-1 \leqq x \leqq +1$.

Diese spezielle, besonders einfache Herstellungsregel für die Bildung Weierstraßscher Approximationspolynome durch Hermitesche Interpolation habe ich im Jahre 1915 veröffentlicht[7]. In der neuen allgemeinen Herstellungsregel im § 11 figurieren nun *beliebige* Neigungswerte $y_1', y_2', \ldots, y_n'$ der Hermiteschen Parabel, die nur der Bedingung der *Beschränktheit* ($|y_k'| \leqq \Delta$ für jedes k und jedes n) unterworfen sein müssen. Nimmt man für diese — um einen anderen speziellen Fall zu erwähnen — den Wert

$$y_1' = y_2' = \ldots = y_n' = 1,$$

so erhält man den Satz, daß unsere Hermite-Parabeln auch dann gleich-

---

[7]) L. Fejér, a) Interpolatióról (ungarisch), Anzeiger der Ungarischen Akademie der Wissenschaften **34** (1916) (Sitzung vom 15. November 1915); b) Über Interpolation, Nachrichten der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse, 1916 (Sitzung am 15. Januar 1916). Mit dieser Arbeit sind die §§ 8—12 vorliegender Arbeit zu vergleichen; jene enthält auch wichtige literarische Verweise.

mäßig zu $f(x)$ konvergieren (immer mit Tschebyscheff-Abszissen!), wenn sämtliche Tangenten dieser Parabeln in den Punkten

$$(x_1, f(x_1)), \quad (x_2, f(x_2)), \quad \ldots, \quad (x_n, f(x_n))$$

45^0 mit der Abszissenachse bilden.

Ich glaube, daß man durch zweckmäßige Wahl der „Tangentenverteilung" in unserem Hermite-Polynome noch gewisse weitere Vorteile wird gewinnen können.

21. Ich erwähne schließlich noch einen dritten, etwas anders gearteten Spezialfall: es habe $f(x)$ selbst im Intervalle $-1 \leqq x \leqq +1$, oder wenigstens für sämtliche Tschebyscheffschen Abszissen, einen *beschränkten* Differentialquotienten $f'(x)$. Dann kann ich $y'_k = f'(x_k)$ wählen. Die Hermitesche Parabel geht in diesem Falle in die *Schmiegungsparabel* $y = S_n(x)$ von $y = f(x)$ über, wo $S_n(x)$ diejenige ganze rationale Funktion von höchstens $(2n-1)$-tem Grade bezeichnet, für die

$$S_n(x_1) = f(x_1), \quad S_n(x_2) = f(x_2), \quad \ldots, \quad S_n(x_n) = f(x_n),$$
$$S'_n(x_1) = f'(x_1), \quad S'_n(x_2) = f'(x_2), \quad \ldots, \quad S'_n(x_n) = f'(x_n)$$

gilt. Unser allgemeiner Satz im § 11 liefert nun den Satz, daß *die Schmiegungsparabel in diesem Falle für $n = \infty$ gleichmäßig zur stetigen Kurve $y = f(x)$ konvergiert* (*d. h.* $\lim S_n(x) = f(x)$, *gleichmäßig für* $-1 \leqq x \leqq +1$).

Budapest, den 28. Juni 1929.

(Eingegangen am 15. 7. 1929.)

Über die Reduzibilität von Polynomen im Körper der reellen Zahlen.

Von

Karl Dörge in Köln.

Um den in dieser Arbeit bewiesenen Satz als Verallgemeinerung einfacher bekannter Theoreme erscheinen zu lassen, knüpfen wir an die Frage nach der Realität der Wurzeln der quadratischen Gleichung an. Der Verallgemeinerung zu Liebe formulieren wir den hier vorliegenden Sachverhalt in folgender etwas umständlicher Form: In $f(x) = a_0 x^2 + a_1 x + a_2$ betrachte man a_0, a_1, a_2 als reelle Parameter. Die Punkte des reellen dreidimensionalen Parameterraumes mit den Koordinaten a_0, a_1, a_2, in welchen $f(x)$ im Körper der reellen Zahlen zerfällt, sind dann durch eine algebraische Ungleichung — nämlich $4a_0 a_2 - a_1^2 \leqq 0$ — charakterisiert. Der Bereich, in welchem f reduzibel ist, wird von dem Bereich, in welchem f irreduzibel ist, getrennt durch die Punktmenge der Punkte, in welchen f das positiv oder negativ genommene Quadrat eines reellen Polynoms wird.

Es sei nun f ein Polynom von den Veränderlichen $x_1, \ldots, x_m$. Die Koeffizienten seien Polynome mit reellen Koeffizienten der reellen Parameter $t_1, \ldots, t_s$. *Unter Q verstehe man die Menge der Punkte des s-dimensionalen Parameterraumes, für welche f oder $-f$ das Quadrat eines reellen Polynoms in $x_1, \ldots, x_m$ wird.* Es ist leicht zu sehen, daß diese Punktmenge Q algebraisch bestimmt ist. Es wird zunächst bewiesen, daß die Punkte des reellen s-dimensionalen Parameterraumes, für welche f im Körper der reellen Zahlen zerfällt, durch Systeme algebraischer Gleichungen und Ungleichungen zwischen den Koeffizienten a_λ von f, also auch zwischen den t_σ, charakterisiert sind. Damit wird dann der folgende Satz bewiesen:

Die Punktmenge Q zerlegt den s-dimensionalen Parameterraum in durch sie getrennte Bereiche. Greift man irgendeinen derartigen Bereich heraus, so ist darin entweder f für jeden Punkt im Reellen reduzibel bis

auf höchstens die Punkte, in denen sich der Grad von f erniedrigt, oder f ist in jedem Punkt des Bereiches irreduzibel bis auf höchstens die Punkte auf gewissen angegebenen algebraischen Flächen[1]).

Bei Polynomen einer Veränderlichen von mindestens drittem Grade z. B. haben alle Gebiete denselben Charakter. In allen ihren Punkten, wo nicht Graderniedrigung eintritt, muß f im Reellen zerfallen. Das Polynom $tx^2+(t-1)y^4$ ist in $0 \leqq t \leqq 1$ reduzibel, in $t<0$ und $t>1$ irreduzibel. Die Grenzpunkte dieser Gebiete sind 0 und 1, wo f ein Quadrat wird.

Aus dem Satze kann z. B. gefolgert werden: Wenn f nie ein Quadrat wird[2]), so ist es entweder im ganzen Raume im Reellen reduzibel bis auf höchstens die Punkte, wo Graderniedrigung eintritt, oder es ist im ganzen Raume irreduzibel bis auf höchstens die Punkte auf angegebenen algebraischen Ausnahmeflächen.

Die Anregung zu der Note habe ich durch das algebraische Kriterium für absolute Irreduzibilität von E. Noether (Math. Annalen 85) bekommen. Ich benutze nicht nur die Sätze, sondern auch die Methoden von E. Noether.

1. $f(x_1, \ldots, x_m)$ sei ein Polynom mit Koeffizienten aus dem Körper K. Der Grad sei h. Man setze $d=2h$. Dann unterwerfe man die x der Transformation $x_\mu = \xi^{d^{m-\mu}}$, $[\mu = 1, 2, \ldots, m]$. f geht dadurch über in ein Polynom der einen Veränderlichen ξ. Man nenne es $F(\xi)$. Die Reduzibilität von f in K läßt sich dann in gewissem Sinne auf die Reduzibilität von F zurückführen[3]). Dazu betrachte man im Körper K die Zerlegung

$$F(\xi) = F_1(\xi) \cdot F_2(\xi).$$

Die in F_1 und in F_2 wirklich auftretenden Exponenten entwickle man als d-adische Zahlen. Wenn die dabei vorkommenden Koeffizienten der Potenzen von d sämtlich ihrem Betrage nach kleiner als $\frac{d}{2}$ sind, sagen wir, die Faktoren F_1 und F_2 hätten induzierte Exponenten. Dann gilt der Satz: f zerfällt dann und nur dann in K, wenn F in zwei Faktoren mit induzierten Exponenten in K zerfällt. Ferner erhält man aus einer Zerlegung von F in Faktoren mit induzierten Exponenten eine Zerlegung von f, in deren Faktoren als Koeffizienten gerade die Koeffizienten der Faktoren von F auftreten.

---

[1]) Hierbei ist der uneigentliche Fall zugelassen, daß der Raum durch Q nicht zerlegt wird, also einen einzigen Bereich bildet. Randpunkte dürfen dem Bereiche angehören.

[2]) Z. B. wenn die bei den verschiedenen Numerierungen der Veränderlichen x_μ sich ergebenden höchsten Glieder und das absolute Glied nirgends sämtlich übereinstimmende Vorzeichen haben, wie z. B. in $x^k - y^k + U(x, y)$, wo U in x und in y höchstens den Grad $k-1$ hat.

[3]) Vgl. die Note von E. Noether.

Im folgenden lege man in den Veränderlichen x irgendeine feste Reihenfolge zugrunde, der Einfachheit halber etwa die Reihenfolge nach den Indizes: $x_1, x_2, \ldots, x_m$. Man kann dann die Glieder von f nach irgendeinem Prinzip, etwa nach fallenden Exponenten, ordnen. Dann schreibe man $f(x_1, \ldots, x_m) = a_0 \Pi_0 + \ldots + a_L \Pi_L = \sum_{\lambda=0}^{L} a_\lambda \Pi_\lambda$, unter $\Pi_0, \ldots, \Pi_L$ die Potenzprodukte der $x_1, \ldots, x_m$ verstanden. Wir setzen im folgenden immer $a_0 \neq 0$ voraus. Der Grad von $F(\xi)$ sei N. Das höchste Glied von F ist dann $a_0 \xi^N$.

Definition. f zerfällt im Körper K halb, wenn es in einem Erweiterungskörper von K in zwei Faktoren $f_1 \cdot f_2$ zerfällt derart, daß in den f_k $[k = 1, 2]$ die höchsten und die konstanten Koeffizienten zu K gehören.

Der Begriff des Halbzerfallens von f läßt sich nun natürlich wieder zurückführen auf einen gewissen Begriff des Halbzerfallens von $F(\xi)$. Dazu sage man:

Definition. F zerfällt in K halb, wenn es in einem Erweiterungskörper von K derart in zwei Faktoren $F_1 \cdot F_2$ zerfällt, daß

1. F_1 und F_2 nur induzierte Exponenten enthalten,

2. die höchsten und die konstanten Koeffizienten von F_1 und F_2 zu K gehören.

Der Bequemlichkeit halber normiere man F, indem man mit a_0^{N-1} multipliziert, dann $\xi' = a_0 \xi$ setzt und schließlich statt ξ' wieder ξ schreibt. Zu $F(\xi)$ gehört dann ein bestimmtes Polynom $F^*(\xi)$, dessen Koeffizienten statt der a_λ jetzt noch mit Potenzen von a_0 multipliziert sind. Offenbar zerfällt F dann und nur dann halb in K, wenn F^* in K halbzerfällt. Ferner zerfällt offenbar f dann und nur dann halb in K, wenn F in K halbzerfällt, also dann und nur dann, wenn F^* in K halbzerfällt.

Wir interessieren uns in Zukunft natürlich nicht für das Halbzerfallen von f, sondern für das Zerfallen von f. Es besteht aber folgender Zusammenhang: Man unterwerfe die Veränderlichen x_μ unter Benutzung irgendwelcher Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_m$ aus K der Transformation $x_\mu = x'_\mu + \alpha_\mu$ $[\mu = 1, 2, \ldots, m]$.

Aus dem durch Division von f mit a_0 normierten Polynom $f^*(x_\mu)$ entsteht dann ein neues Polynom $\hat{f}^*(x'_\mu)$. Wenn f^* in dem[4]) algebraisch abgeschlossenen Erweiterungskörper von K zerfällt, so betrachte man dort sämtliche Zerlegungen von f^* in normierte Faktoren. Aus diesen Zer-

[4]) Unter dem algebraisch abgeschlossenen Erweiterungskörper von K verstehen wir etwa den kleinsten algebraisch abgeschlossenen Erweiterungskörper von K.

legungen erhält man offenbar durch die Transformation $x_\mu = x'_\mu + \alpha_\mu$ sämtliche Zerlegungen von $\hat{f}^*(x'_\mu)$ in normierte Faktoren. Gehört so zu einer Zerlegung $f^* = f_1^* \cdot f_2^*$ die Zerlegung $\hat{f}^* = \hat{f}_1^* \cdot \hat{f}_2^*$, so ist offenbar der absolute Koeffizient von $\hat{f}_1^*$ der Wert $f_1^*(\alpha_\mu)$ und der absolute Koeffizient von $\hat{f}_2^*$ ist $f_2^*(\alpha_\mu)$. Aus einem Zerfallen von f^* in K folgt mithin gewiß das Halbzerfallen von $\hat{f}^*(x'_\mu)$ in K. Umgekehrt folgt aus dem Halbzerfallen von $\hat{f}^*(x'_\mu)$ in K in die Faktoren $\hat{f}_1^* \cdot \hat{f}_2^*$, daß $f_1^*(\alpha_\mu)$ und $f_2^*(\alpha_\mu)$ zu K gehören.

Daraus aber ergibt sich: f^* zerfällt dann und nur dann in K, wenn $f^*(x_\mu + \alpha_\mu)$ für genügend viele und genügend verteilte Systeme α_μ aus K in K halbzerfällt, und zwar ist notwendig und hinreichend, daß nach Bestimmung einer genügend großen Anzahl S, die allein durch die Grade von f bestimmt werden kann, für die Systeme $\alpha_1, \ldots, \alpha_m$, in welchen α_μ unabhängig voneinander die Reihe $0, 1, \ldots, S-1$ durchlaufen, $f^*(x_\mu + \alpha_\mu)$ in K halbzerfällt[5]). Kehrt man zu den Polynomen vor der Normierung zurück, so ergibt sich: f zerfällt dann und nur dann in K, wenn nach Bestimmung einer hinreichend großen Zahl S, die allein durch die Grade von f relativ zu den m Veränderlichen x_μ bestimmt werden kann, $f(x_\mu + \alpha_\mu)$ für alle S^m Systeme von Zahlen α_μ, in welchen die α_μ unabhängig voneinander ganzzahlig von 0 bis $S-1$ variieren, als Polynom von den x_μ in K halbzerfällt.

2. Um eine Bedingung für das Halbzerfallen von f in K zu erhalten, genügt es, das Halbzerfallen von F^* in K zu betrachten. Dazu nehme man für irgendein festes ϱ der Reihe $1, 2, \ldots, N-1$ irgendeine Zerlegung von F^* in einem Erweiterungskörper von K, wobei man die Faktoren als normiert annehme:

$$F^* = F_1^* \cdot F_2^*,$$

$$F_1^* = \xi^\varrho + c_1 \xi^{\varrho-1} + \ldots + c_\varrho, \quad F_2^* = \xi^{N-\varrho} + e_1 \xi^{N-\varrho-1} + \ldots + e_{N-\varrho}.$$

Die N Wurzeln von F seien $\xi_1, \xi_2, \ldots, \xi_N$. Sie verteilen sich auf die Faktoren F_1 und F_2. Man wähle die Bezeichnung etwa so, daß F_1 die Wurzeln $\xi_1, \ldots, \xi_\varrho$ und F_2 die Wurzeln $\xi_{\varrho+1}, \ldots, \xi_N$ hat. Daß diese Zerlegung von F^* dann ein Halbzerfallen von F^* liefert, bedeutet: Gewisse der c, nämlich die, welche zu nicht induzierten Exponenten gehören, etwa $c_{\alpha_1}, \ldots, c_{\alpha_\varkappa}$, und ebenso gewisse der e, etwa $e_{\beta_1}, \ldots, e_{\beta_\sigma}$ sollen verschwinden und außerdem soll c_ϱ — und dann von selbst $e_{N-\varrho}$ — zu K gehören. Das kann aber auch so ausgedrückt werden: Der mit den neuen Veränderlichen $u_1, u_2, \ldots, v_1, v_2, \ldots, w, \delta$ gebildete Ausdruck

$$u_1 c_{\alpha_1} + u_2 c_{\alpha_2} + \ldots + u_k c_{\alpha_\varkappa} + v_1 e_{\beta_1} + v_2 e_{\beta_2} + \ldots + v_\sigma e_{\beta_\sigma} + w(c_\varrho - \delta)$$

[5]) Vgl. den Hilfssatz in Teil II der Note „Bemerkung zum Hilbertschen Irreduzibilitätssatz", Math. Annalen 102, S. 521–530.

soll für eine Zahl δ aus K verschwinden. Auf diese Form, aufgefaßt als Ausdruck in den $\xi_1, \ldots, \xi_N$ wende man alle $N!$ Permutationen der $\xi_1, \ldots, \xi_N$ an. Man erhält dann weitere Formen in denselben Veränderlichen u, v, w, δ. Man bilde deren Produkt. Dieses ist ein Polynom $\Phi_\varrho(u, v, w, \delta)$ von u, v, w, δ, dessen Koeffizienten jetzt wegen ihrer Symmetrie in den $\xi_1, \ldots, \xi_N$ Polynome von den Koeffizienten von F^*, also auch von den Koeffizienten a_λ mit ganzen rationalen Koeffizienten sind. Zu jedem ϱ der Reihe $1, 2, \ldots, N-1$ gehört also ein $\Phi_\varrho(u, v, w, \delta)$. Man nehme hier der Einfachheit halber für die verschiedenen ϱ etwa dieselben Unbestimmten u, v, w, δ, deren Anzahl man von vornherein hinreichend groß annehme. Die Bedingung für das Halbzerfallen von f in K ist dann: $\Phi(\delta) = 0$ soll bei veränderlichen u, v, w eine Lösung δ in K haben.

Setzt man $w = 0$, so entsteht aus Φ ein Polynom Ψ, welches nur noch von u, v abhängt. Das Verschwinden von Ψ ist dann offenbar die notwendige und hinreichende Bedingung für das Zerfallen von f in dem algebraisch abgeschlossenen Erweiterungskörper von K.

In ähnlicher Weise kann man eine Form Ψ_3 von Unbestimmten $u_1, u_2, \ldots, v_1, v_2, \ldots, p_1, p_2, \ldots$ konstruieren mit Koeffizienten, welche Polynome der a_λ mit ganzen rationalen Koeffizienten sind derart, daß $\Psi_3 = 0$ die notwendige und hinreichende Bedingung dafür ist, daß f in dem algebraisch abgeschlossenen Erweiterungskörper von K in mindestens drei Faktoren zerfällt.

Durch die S^m Transformationen $x_\mu = x'_\mu + \alpha_\mu$, $\alpha_\mu = 0, 1, \ldots, S-1$ enstehen aus f S^m Polynome f_σ. Zu jedem f_σ gehört ein bestimmtes $\Phi(\delta)$, man nenne es $\Phi_\sigma(\delta)$. Notwendig und hinreichend für das Zerfallen von f in K ist dann: Jedes Φ_σ $[\sigma = 1, 2, \ldots, S^m]$ soll in K eine Wurzel δ haben.

3. An Stelle des beliebigen Körpers K betrachte man jetzt den Körper aller reellen Zahlen, R. Dann wähle man ein festes σ. Die Koeffizienten von Φ_σ, aufgefaßt als Polynom der u, v, w, sind Polynome von δ mit reellen Koeffizienten. Man bezeichne sie etwa mit $P_k(\delta)$ $[k = 1, 2, \ldots, K]$ und setze $\sum_{k=1}^{K} (P_k(\delta))^2 = Q_\sigma(\delta)$. Die Bedingung für das Zerfallen von f in R ist dann die: Jedes $Q_\sigma(\delta)$ $[\sigma = 1, 2, \ldots, S^m]$ soll eine reelle Wurzel δ haben.

Ob bei festem σ das Polynom $Q_\sigma(\delta)$ eine reelle Wurzel hat, ist nun durch Gleichungen und Ungleichungen zwischen den Koeffizienten von Q_σ bestimmt, wie aus den bekannten Sätzen über die Anzahl der reellen Wurzeln von Polynomen folgt. Genauer gilt folgendes: Man kann zu $Q_\sigma(\delta)$ eine endliche Anzahl von Systemen $\Sigma_{\sigma\tau}$ $[\tau = 1, \ldots, \tau_\sigma]$ bestimmen, von denen jedes aus einer endlichen Anzahl von algebraischen Gleichungen und algebraischen

Ungleichungen*) zwischen den Koeffizienten von Φ_σ, also zwischen den Koeffizienten von f besteht, mit folgender Eigenschaft: Dafür, daß $Q_\sigma(\delta)$ eine reelle Wurzel hat, ist notwendig und hinreichend, daß für irgendein τ alle Gleichungen und Ungleichungen von $\Sigma_{\sigma\tau}$ gleichzeitig erfüllt sind. Daraus ergibt sich weiter: Notwendig und hinreichend für das Zerfallen von f in R ist, daß für jedes σ ein System $\Sigma_{\sigma\tau}$ existiert, dessen Gleichungen und Ungleichungen gleichzeitig erfüllt sind.

Offenbar kann man dann auch ohne Unterscheidung der σ von vornherein endlich viele Systeme P_ζ von je endlich vielen algebraischen Relationen zwischen den Koeffizienten von f angeben, so daß für das Zerfallen von f in R notwendig und hinreichend ist, daß für irgendein ζ sämtliche Relationen von P_ζ gleichzeitig erfüllt sind.

Diese Bedingung hat sich ergeben unter der — allerdings einzigen — Einschränkung, die wir bisher durchweg gemacht haben, nämlich daß a_0, der Koeffizient des höchsten Gliedes von f, nicht verschwindet. Ein Satz gleichen Wortlautes ergibt sich nun nachträglich auch, wenn man diese Einschränkung fallen läßt.

Statt P_ζ schreibe man nämlich jetzt $P^0_{\zeta_0}$ $(\zeta_0 = 1, 2, \ldots, Z_0)$. Entsprechend erhält man, wenn man $a_0 = 0$, $a_1 \neq 0$ annimmt, ein System algebraischer Relationen $P^1_{\zeta_1}$ $[\zeta_1 = 1, \ldots, Z_1]$ gleicher Bedeutung, für $a_0 = 0$, $a_1 = 0$, $a_2 \neq 0$ ein System $P^2_{\zeta_2}$ usw. Vor jede Relation von $P^0_{\zeta_0}$ schreibe man nun noch formal $a_0^2 > 0$, was durch $\begin{pmatrix} a_0^2 > 0 \\ P^0_{\zeta_0} \end{pmatrix}$ angedeutet werde. Man erhält dann ein System von algebraischen Relationen, das mit $S^0_{\zeta_0}$ bezeichnet werden möge.

Vor jede Relation von $P^1_{\zeta_1}$ setze man die beiden Relationen $a_0 = 0$, $a_1^2 > 0$, was durch $\begin{pmatrix} a_0 = 0 \\ a_1^2 > 0 \\ P^1_{\zeta_1} \end{pmatrix}$ angedeutet werde. Das System von Relationen, das man so erhält, nenne man $S^1_{\zeta_1}$. Entsprechend erhält man $S^2_{\zeta_2}$: $\begin{pmatrix} a_0 = 0 \\ a_1 = 0 \\ a_2^2 > 0 \\ P^2_{\zeta_2} \end{pmatrix}$ usw. Mithin gilt

Satz I: *Notwendig und hinreichend für das Zerfallen von f in R ist, daß irgendeines der endlich vielen* $Z_0 + Z_1 + \ldots$ Systeme S_ϑ, $\vartheta = 1, 2, \ldots, H$, $[Z_0 + Z_1 + \ldots = H]$ *von algebraischen Gleichungen und Ungleichungen zwischen den Koeffizienten von f gleichzeitig erfüllt ist.*

*) Unter einer algebraischen Ungleichung verstehen wir eine Behauptung, daß ein Polynom positiv oder negativ oder nicht positiv oder nicht negativ ist.

Unter einer algebraischen Relation verstehen wir eine algebraische Gleichung oder eine algebraische Ungleichung.

Es liegt also prinzipiell gerade so, wie bei der Frage nach der Existenz reeller Wurzeln von Polynomen mit einer einzigen Unbekannten, natürlich rechnerisch viel komplizierter.

Im $L+1$-dimensionalen Raume der reellen Parameter $a_0, \ldots, a_L$ gibt es also endlich viele algebraische Flächen, nämlich die, welche man erhält, wenn man die Relationen, die in irgendeinem der H Systeme S_ϱ vorkommen, durch Ersetzen der etwa vorkommenden Ungleichheitszeichen durch Gleichheitszeichen sämtlich zu Gleichungen macht, derart, daß man von einem Punkt, in welchem f reduzibel ist, zu einem Punkt, in welchem f irreduzibel ist, nur gelangen kann durch Passieren einer derselben.

Wir werden jetzt diese Flächen genauer zu charakterisieren suchen.

4. Statt die Koeffizienten a_λ selbst als unabhängige Veränderliche anzunehmen, nehmen wir allgemeiner an, sie seien Polynome der unabhängigen Parameter $t_1, \ldots, t_s$ mit reellen Koeffizienten. Nun behandeln wir zunächst den Fall eines einzigen Parameters t. Den allgemeinen Fall führen wir auf diesen zurück.

Es mögen also die Koeffizienten a_λ Polynome mit reellen Koeffizienten von t sein, $a_0(t) \neq 0$.

Aus den algebraischen Flächen in den a_λ werden jetzt Polynome in t. Sie verschwinden nicht sämtlich identisch in t, weil $(a_0(t))^2$ vorkommt. Die höchstens endlich vielen reellen Nullstellen der nicht identisch verschwindenden unter ihnen teilen offenbar die reelle t-Achse ein in endlich viele Strecken, so daß jede von ihnen entweder nur Punkte enthält, wo f in R reduzibel ist, oder nur Punkte, wo f in R irreduzibel ist. Es kann sein, daß zu dieser Einteilung gar nicht alle Punkte, die man erhalten hat, erforderlich sind. Dann streiche man sie fort. Die übrigen Punkte nenne man die Grenzpunkte der Einteilung. Sie sind durch f eindeutig bestimmt.

$A(t)$ sei der Körper aller algebraischen Funktionen von t. Er ist algebraisch abgeschlossen. Macht man die Betrachtung von 1. und 2. statt mit dem beliebigen K jetzt mit $A(t)$, so ist mithin $\Psi(u, v) = 0$ die Bedingung dafür, daß f in $A(t)$ in mindestens zwei Faktoren zerfällt, $\Psi_3(u, v, p) = 0$ die Bedingung dafür, daß es in mindestens drei Faktoren zerfällt. Bei reell oder komplex spezialisiertem t ist ferner für solche t, für welche $a_0(t) \neq 0$ ist, $\Psi(u, v) = 0$ und $\Psi_3(u, v, p) = 0$ die Bedingung dafür, daß f im Körper $R(i)$ aller reellen und komplexen Zahlen in mindestens zwei bzw. drei Faktoren zerfällt. Dann unterscheiden wir bzgl. $A(t)$ drei Fälle:

$$\text{(I)} \qquad \Psi(u, v) \neq 0.$$

Dann gibt es höchstens endlich viele reelle Punkte t, an welchen Ψ verschwindet. Außerdem gibt es nur endlich viele Stellen, an welchen $a_0(t)$

verschwindet. Höchstens an diesen endlich vielen Stellen, an welchen Ψ oder a_0 verschwindet, kann f in $R(i)$, also erst recht nur hier in R zerfallen. Wir haben dies geschlossen, indem wir eine bestimmte Reihenfolge der Veränderlichen x_μ ausgezeichnet hatten. Zu dem entsprechenden Ergebnis hätten wir mit jeder anderen Reihenfolge gelangen müssen. Daher gilt der Satz: Ist $\Psi(u,v) \neq 0$, so kann f höchstens an den endlich vielen reellen Stellen t zerfallen, an welchen $\Psi(u,v)$ verschwindet, oder bei jeder der $m!$ möglichen zugrunde gelegten Reihenfolgen $x_{a_1}, \ldots, x_{a_m}$ der Veränderlichen das Leitglied herausfällt, d. h. aber, an welchen Graderniedrigung eintritt. Das Verschwinden von $\Psi(u,v)$ für reelles t ist gleichbedeutend mit dem Verschwinden des Polynoms $\hat{\Psi}(a_\lambda)$, welches entsteht, indem man die Summe der Quadrate aller Koeffizienten von $\Psi(u,v)$ bildet.

Es sei jetzt

$$\text{(II)} \qquad \Psi(u,v) = 0, \quad \Psi_3(u,v,p) = 0.$$

Dann enthält f an jeder reellen Stelle t, an welcher $a_0(t) \neq 0$ ist, in $R(i)$ mindestens drei Faktoren. Da neben jedem komplexen Faktor der konjugiert komplexe Faktor auftritt, muß dann an jeder reellen Stelle t f auch in R zerfallen bis auf höchstens die Stellen, an welchen $a_0(t)$ verschwindet.

Legt man eine andere Reihenfolge der Veränderlichen x_μ zugrunde, so erhält man ein neues Ψ und ein neues Ψ_3. Offenbar müssen aber auch diese beiden verschwinden. Daher kann man allgemein schließen: f zerfällt an allen reellen Punkten, außer höchstens an denjenigen, wo Graderniedrigung eintritt.

$$\text{(III)} \qquad \Psi(u,v) = 0, \quad \Psi_3(u,v,p) \neq 0.$$

Diejenigen der endlich vielen reellen Punkte t, an welchen $a_0(t)$ verschwindet, bezeichne man mit $\tau_1, \tau_2, \ldots$, die, an welchen Ψ_3 verschwindet, mit $\tau_1', \tau_2', \ldots$. In jedem reellen Punkte, der kein τ ist, enthält dann f in $R(i)$ mindestens zwei Faktoren; wenn der Punkt auch kein τ' ist, genau zwei Faktoren.

f zerfällt nun in $A(t)$ in zwei Faktoren:

$$\text{(x)} \qquad f(x_\mu) = (\varphi_0 \Pi_0 + \ldots + \varphi_K \Pi_K) \cdot (\psi_0 \Pi_0' + \ldots + \psi_{K'} \Pi_{K'}').$$

Durch Spezialisieren von t in den φ und ψ erhält man aus dieser Zerlegung Zerlegungen von f bei numerischem t. Da f bis auf die endlich vielen τ, τ' genau zwei Faktoren enthält, folgt, daß man bei eindeutiger Festlegung eines Zweiges der φ, ψ, derart daß die Zerlegung (x) stattfindet, bereits sämtliche Zerlegungen von f an allen Stellen außer den τ, τ' bekommt. Wir nehmen also in Zukunft bei numerischer Betrachtung an,

daß die φ, ψ bei anzugebenden Normierungen von φ_0 und ψ_0 eindeutige Zweige der durch sie dargestellten algebraischen Funktionen sind.

Wir sehen uns nun die endlich vielen Grenzpunkte t, durch die wir die ganze reelle t-Achse eingeteilt haben, einzeln an:

III. 1. t_0 sei Grenzpunkt zwischen zwei reduziblen Strecken [7]. Die φ und ψ normiere man so, daß $\varphi_0 = 1$ ist. Die φ und ψ müssen dann in einer ganzen Umgebung von t_0 außer höchstens an der Stelle t_0 selbst reelle Werte haben. Wenn t_0 nun kein τ ist, sind die φ und ψ an der Stelle t_0 stetig. Daher ist an der Stelle t_0:

$$f_{t_0} = [(\lim \varphi_0)\,\Pi_0 + \ldots + (\lim \varphi_K)\,\Pi_K] \cdot [\lim \psi_0\,\Pi_0' + \ldots + (\lim \psi_{K'})\,\Pi_{K'}']\,.$$

Wegen $a_0(t_0) \neq 0$ ist hier $\lim \varphi_0 \cdot \lim \psi_0 \neq 0$, ferner sind sämtliche Limites reell, also zerfällt f in R. Ein Grenzpunkt zwischen reduziblen Strecken muß mithin ein τ sein, a_0 muß also dort verschwinden; also da man dies wieder für jede Reihenfolge der x_u schließen kann, muß Graderniedrigung eintreten.

III. 2. t_0 sei Grenzpunkt zwischen einer irreduziblen und einer reduziblen Strecke. Man operiere nun in so kleinen Umgebungen u um t_0, daß darin außer etwa t_0 kein weiteres τ, τ' vorkommt. $\mathfrak{J}$ und $\mathfrak{R}$ seien die Teile von u, welche in die irreduzible bzw. die reduzible Strecke hineinfallen. In $\mathfrak{J}$ kann a_0 dann nicht das Vorzeichen wechseln. Dann normiere man φ_0 so, daß in $\mathfrak{J}$ gilt:

$$\text{wenn in } \mathfrak{J}\ a_0 > 0: \quad \varphi_0 = \sqrt{a_0(t)}, \quad \psi_0 = \sqrt{a_0(t)},$$
$$\text{wenn in } \mathfrak{J}\ a_0 < 0: \quad \varphi_0 = i\sqrt{-a_0(t)}, \quad \psi_0 = i\sqrt{-a_0(t)}.$$

Offenbar sind dann die φ und ψ in $\mathfrak{J}$ konjugiert komplex, wenn $a_0 > 0$, sie sind bis auf den Faktor -1 konjugiert komplex, wenn $a_0 < 0$. In $\mathfrak{R}$ sind die φ bis auf einen etwaigen gemeinsamen komplexen Faktor reell, ebenso die ψ. Ferner sind bei dieser Normierung, da t_0, wie aus der Betrachtung von $\mathfrak{J}$ folgt, keine Unendlichkeitsstelle der φ und φ ist, die φ und ψ an der Stelle t_0 stetig. Daher gilt an der Stelle t_0 die Zerlegung

$$f_{t_0} = [(\lim \varphi_0)\,\Pi_0 + \ldots + (\lim \varphi_K)\,\Pi_K] \cdot [(\lim \psi_0)\,\Pi_0' + \ldots + (\lim \psi_{K'})\,\Pi_{K'}']\,.$$

Nimmt man diese Limites zunächst bei Annäherung an t_0 aus $\mathfrak{J}$ heraus, so folgt, daß die Faktoren entweder konjugiert komplex sind, oder daß der eine Faktor das -1-fache des konjugiert komplexen des anderen Faktors ist, also an der Stelle t_0 gilt entweder

$$f = f_1 \bar{f}_1 \qquad \text{oder} \qquad f = \bar{f}_1(-\bar{f}_1)\,.$$

[7] Wir nennen eine Strecke kurz reduzibel, wenn f in jedem ihrer Punkte reduzibel ist. Entsprechend nennen wir sie irreduzibel, wenn f in ihr überall irreduzibel ist.

Nimmt man aber die Limites durch Annäherung von t_0 aus $\mathfrak{R}$ heraus, so folgt, daß an der Stelle t_0 f_1 und $\bar{f}_1$ durch Herausnahme zweier konstanter Faktoren C_1, C_2 aus f_1 bzw. $\bar{f}_1$ reell werden. Man kann dasselbe dann auch erreichen, indem man $C_2 = \bar{C}_1$ wählt, da $C_1 C_2$ reell sein muß. $\frac{f_1}{C_1}$ und $\frac{\pm f_1}{C_2}$ sind dann aber reell und bis auf das Vorzeichen konjugiert komplex, also reell und bis auf das Vorzeichen gleich, etwa gleich $\pm f_3$. Dann hat man an der Stelle t_0 eine Zerlegung

$$f = C_1 \bar{C}_1 f_3^2 \qquad \text{oder} \qquad f = C_1 \bar{C}_1 (-f_3^2),$$

d. h.

$$f = (\sqrt{C_1 \bar{C}_1}\, f_3)^2 \qquad \text{oder} \qquad f = -(\sqrt{C_1 \bar{C}_1}\, f_3)^2.$$

Mithin ist f an der Stelle t_0 das Quadrat oder das negativ genommene Quadrat eines reellen Polynoms oder, was dasselbe ist, das Quadrat eines Polynoms[8]).

III. 3. Schließlich sei t_0 Grenzpunkt zwischen zwei irreduziblen Strecken. Mithin zerfällt f an der Stelle t_0 in reelle Faktoren. Daraus schließt man genau wie in III. 2, daß f an der Stelle t_0 in zwei konjugiert komplexe Faktoren, multipliziert mit ± 1 zerfällt. Wenn f dann an der Stelle t_0 nur in zwei Faktoren zerfällt, ergibt sich genau wie in III. 2, daß an der Stelle t_0 f das Quadrat eines Polynoms ist. Daher erhält man: Entweder ist f bei t_0 ein Quadrat oder f zerfällt an der Stelle t_0 in mindestens drei Faktoren. Dafür ist, wenn $a_0(t_0) \neq 0$, $\Psi_3 = 0$ notwendig und hinreichend. Mit Ψ_3 kann man wieder wie in 4. I zur Quadratsumme der Koeffizienten von Ψ_3 übergehen.

Wenn man die Betrachtungen aus 2. für algebraisch abgeschlossene Körper, wie es E. Noether tut, mit homogenisierten Polynomen durchführt, so erhält man allgemeiner statt unseres Ψ_3 und $\hat{\Psi}_3$ ein Ψ_3^* und ein $\hat{\Psi}_3^*$, dessen Verschwinden notwendig und hinreichend dafür ist, daß f in $R(i)$ mindestens drei Faktoren enthält, außer wenn Graderniedrigung eintritt. Man erhält dann etwas schärfer, daß f bei t_0 entweder ein Quadrat ist oder Graderniedrigung eintritt, oder $\hat{\Psi}_3^*$ verschwindet. Dabei ist $\hat{\Psi}_3 = 0$ bei nicht eintretender Graderniedrigung für das Zerfallen in R notwendig und hinreichend.

5. Hängt f nicht nur von einem Parameter t, sondern von s Parameter $t_1, \ldots, t_s$ ab, so gelten die abgeleiteten Sätze zunächst nur, wenn man alle bis auf einen Parameter festhält. Daraus folgt aber sofort zusammengefaßt folgendes allgemeine Resultat: $f(x_1 \ldots x_m)$ sei ein Polynom von

---

[8]) Hierbei muß — und diese Festsetzung treffe man auch für das Folgende — das Polynom f an solchen Stellen, an denen es sich auf eine Konstante reduziert, als ein Quadrat dann und nur dann angesehen werden, wenn es in jeder Umgebung der Stelle das Quadrat eines — nicht konstanten — Polynoms ist.

$x_1, \ldots, x_m$. Die Koeffizienten seien Polynome von s reellen Parmetern $t_1, \ldots, t_s$. Geordnet nach fallenden Potenzen für irgendeine feste Reihenfolge der x_μ sei

$$f = a_0 \Pi_0 \ldots + a_L \Pi_L = \sum_{\lambda=0}^{L} a_\lambda \Pi_\lambda, \qquad a_0 \neq 0.$$

Die Punkte des reellen s-dimensionalen t-Raumes, in welchen f das Quadrat eines Polynoms wird — f ist dann von selbst das ± 1-fache des Quadrates eines reellen Polynoms — bilden die Punktmenge Q, die, worauf am Schluß hingewiesen wird, im wesentlichen algebraisch bestimmt ist. Die Punkte, in welchen f bei nicht verschwindendem a_0 mindestens drei komplexe Teiler enthält, bilden eine algebraische Fläche $\hat{\Phi}_3(a_\lambda) = 0$. Die Punkte, in denen Graderniedrigung in f eintritt, sind selbstverständlich algebraisch bestimmt, denn die Quadratsumme von gewissen a_λ soll verschwinden. Man nenne diese Quadratsumme G und erhält die Fläche $G = 0$. Schließlich brauchen wir noch die algebraische Fläche $\hat{\Phi} = 0$, deren Konstruktion oben angegeben ist, und die bei nicht verschwindendem a_0 angibt, daß f in $R(i)$ zerfällt. Dann gilt folgender

Satz II: *Es sind nur folgende drei Fälle für das Zerfallen in R möglich:*

A) *f ist im ganzen Raume bis auf eine höchstens $s-1$-dimensionale*[9]) *Punktmenge reduzibel. Dann ist f höchstens für die Punkte der Fläche $G = 0$ irreduzibel.*

B) *f ist im ganzen Raume bis auf eine höchstens $s-1$-dimensionale Punktmenge irreduzibel. Dann ist f, wenn $\hat{\Phi} \not\equiv 0$ in den $t_1 \ldots t_s$, höchstens an den Punkten der Fläche $\hat{\Phi} = 0$, und wenn $\hat{\Phi} \equiv 0$ höchstens an den Punkten der Fläche $a_0 = 0$ oder $\hat{\Phi}_3 = 0$ und den Punkten der Punktmenge Q reduzibel. Reduzibilität tritt wirklich ein an den Punkten von Q; an den Punkten von $\hat{\Phi}_3 = 0$, wenn $a_0 \neq 0$.*

C) *f sei für eine s-dimensionale Punktmenge reduzibel und für eine s-dimensionale Punktmenge irreduzibel. Die Punktmenge Q zerlegt dann den ganzen Raum in durch sie getrennte Bereiche derart, daß für jeden einzelnen derselben folgendes gilt:*

Entweder ist f in allen Punkten des Bereiches reduzibel bis auf höchstens die Punkte $G = 0$ oder es ist in allen Punkten des Bereiches irreduzibel bis auf höchstens die Punkte von $a_0 = 0$ und $\hat{\Phi}_3 = 0$. Reduzibilität tritt wirklich ein, wenn $a_0 \neq 0$, aber *$\Phi_3 = 0$. In den Punkten von Q, die diese Bereiche begrenzen, tritt wirklich Reduzibilität ein.*

---

[9]) Eine Punktmenge im $t_1, \ldots, t_s$-Raume heißt höchstens $s-1$-dimensional, wenn sie nur aus Randpunkten besteht. Sie heißt s-dimensional, wenn sie einen Punkt und eine ganze Umgebung desselben enthält.

Zusatz: Statt $\hat{\Phi}_3 = 0$ gibt es eine Fläche $\hat{\Phi}_3^* = 0$, so daß man an Stelle des Paares von Gleichungen $a_0 = 0$, $\hat{\Phi}_3 = 0$ in B) und C) auch die Gleichungen $G = 0$, $\hat{\Phi}_3^* = 0$ mit sonst gleichem Wortlaute hätte schreiben können.

Schließlich werde noch eine Bemerkung über die algebraische Bestimmtheit der Punktmenge Q angeknüpft. Es ist klar, daß bei nichtverschwindendem a_0 das Polynom f dann und nur dann ein Quadrat ist, wenn F in zwei gleiche Faktoren mit induzierten Exponenten zerfällt. Dazu müssen also gewisse der c und ε verschwinden, die übrigen aber, soweit sie gleiche Indizes haben, übereinstimmen. Es ist klar, daß diese Bedingungen entsprechend den Schlüssen von 2. auf algebraische Gleichungen führen.

Durch Benutzung der E. Noetherschen Methode, welche mit homogenisierten Polynomen operiert, erhält man ein System algebraischer Gleichungen, welches nur versagt, wenn Graderniedrigung eintritt.

Schließlich werde noch folgender leicht zu beweisender Zusatz ohne Beweis angegeben.

2. Zusatz: Es gibt, wie soeben angedeutet, eine algebraische Gleichung $Q(a_\lambda) = 0$, die bei nichtverschwindendem a_0 die Bedingung dafür ist, daß f ein Quadrat ist. Unter Q_0 verstehe man die Gesamtheit der Punkte, in denen bei nichtverschwindendem a_0 das Polynom $Q(a_\lambda)$ verschwindet, und die Gesamtheit aller Häufungspunkte derselben. Ferner nehme man zu Q_0 die Punkte der Fläche $a_0 = 0$ dann und nur dann hinzu, wenn f nach Fortlassen des höchsten Gliedes identisch in den t_σ das Quadrat eines Polynoms von den x_μ wird. Diese Punktmenge Q_0 — im allgemeinen ein echter Teil von Q — reicht aus, um die Bereiche verschiedenen Charakters, von denen in dem einen f bis auf algebraische Ausnahmepunkte reduzibel, in dem anderen bis auf algebraische Ausnahmepunkte irreduzibel ist, voneinander zu trennen. An den übrigen Punkten in Q, die mithin nicht zu Q_0 gehören, muß f zerfallen, aber es können dort nicht Gebiete mit verschiedenem Charakter aneinander stoßen.

(Eingegangen am 8. 4. 1929.)

Eine Bemerkung über die Unzerlegbarkeit von Polynomen.

Von

Bartel L. van der Waerden in Groningen (Niederlande).

Aus den Behauptungen auf S. 744 der Abhandlung von G. Hermann über die Frage der endlichvielen Schritte in der Theorie der Polynomideale[1]) könnte man entnehmen, daß es nicht nur bei endlichen, sondern auch bei unendlichen Erweiterungen von (sagen wir etwa) dem Körper Γ der rationalen Zahlen möglich sei, zu entscheiden, ob ein Polynom $f(x)$ im Körper irreduzibel sei. Beim Beweis wird aber eine petitio principii benutzt, nämlich, daß es möglich sei, von jedem endlichen Erweiterungskörper Λ von Γ zu entscheiden, ob im vorgelegten unendlichen Körper ein zu Λ äquivalenter Unterkörper vorhanden ist. Wie man das aber in endlichvielen Schritten entscheiden soll, wird nicht gesagt.

Aus dem Folgenden wird sich ergeben, daß ein allgemeines Irreduzibilitätskriterium für Polynome $f(x)$ in unendlichen, aber „explizite-bekannten" (in einem später zu erklärenden Sinn) Erweiterungskörpern des rationalen Zahlkörpers Γ *unmöglich* ist, es sei denn, daß es der Beweistheorie gelingen wird, die Entscheidbarkeit einer jeden Frage von der Art „Gibt es ein n mit der Eigenschaft $E(n)$?" darzutun. $E(n)$ soll eine Eigenschaft der natürlichen Zahl n sein, deren Gültigkeit für jedes n in endlichvielen Schritten feststellbar ist.

Ein Körper K soll *explizite-bekannt* heißen, wenn seine Elemente Symbole aus einem bekannten abzählbaren Vorrat von unterscheidbaren Symbolen sind, deren Addition, Multiplikation, Subtraktion und Division sich in endlichvielen Schritten ausführen lassen. „Explizite-bekannt" ist z. B. ein Körper $\Gamma(\theta_1, \theta_2, \ldots)$, wo jedes θ_ν entweder eine Unbestimmte oder eine Wurzel einer bestimmten in $\Gamma(\theta_1, \ldots, \theta_{\nu-1})$ irreduziblen Gleichung ist.

[1]) Math. Annalen **95** (1926), S. 736–788.

Behauptung. *Solange man keine allgemeine Methode hat, jedes Problem von der Art „Gibt es ein n mit der Eigenschaft $E(n)$?“ zu lösen, solange kann es auch keine allgemeine Methode der Faktorzerlegung von Polynomen $f(x)$ mit Koeffizienten aus einem explizite-bekannten Körper geben.*

Beweis[2]). Gesetzt, es gäbe eine allgemeine Methode, jedes Polynom mit Koeffizienten aus einem Körper $\Gamma(\theta_1, \theta_2, \ldots)$ in Faktoren zu zerlegen; dann ist zu zeigen, daß man auch jede Frage „Gibt es ein n mit der Eigenschaft $E(n)$?“ lösen kann. Es sei eine Eigenschaft $E(n)$ vorgelegt. Wir konstruieren den Körper $\Gamma(\theta_1, \theta_2, \ldots)$, wo Γ der Körper der rationalen Zahlen ist,

$\theta_n = \sqrt{-1}$, wenn n die kleinste Zahl mit der Eigenschaft $E(n)$,

$\theta_n = \sqrt{p_n}$ sonst, wo p_n die n-te Primzahl ist.

Daß dieser Körper explizite-bekannt ist, ist unmittelbar einzusehen. Ebenso, daß das Polynom $x^2 + 1$ in ihm unzerlegbar ist oder nicht, je nachdem es ein n mit der Eigenschaft $E(n)$ gibt. Die nach Voraussetzung mögliche Entscheidung über die Irreduzibilität des Polynoms $x^2 + 1$ gibt also zugleich Aufschluß über die Existenz eines n mit der Eigenschaft $E(n)$.

Daß die gemachte Voraussetzung wesentlich ist, sieht man so: Wäre das allgemeine Entscheidungsproblem für Fragen der Art „Gibt es ein n mit der Eigenschaft $E(n)$?“ gelöst, so würde man wenigstens für die Körper $\Gamma(\theta_1, \theta_2, \ldots)$ das Irreduzibilitätsproblem ohne weiteres lösen können, denn ein Polynom $f(x)$ ist dann und nur dann in $\Gamma(\theta_1, \theta_2, \ldots)$ zerlegbar, wenn es ein n gibt mit der Eigenschaft, daß $f(x)$ im Körper $\Gamma(\theta_1, \ldots, \theta_n)$ zerlegbar ist, und diese Eigenschaft $E(n)$ ist für jedes n nach der Kroneckerschen Methode (siehe G. Hermann, a. a. O.) entscheidbar.

---

[2]) Bei der Bildung des Beispiels, das den wesentlichen Inhalt dieses Beweises ausmacht, hätte ich mich natürlich auch auf eine bestimmte Eigenschaft $E(n)$, etwa ein bestimmtes bis jetzt noch nicht beobachtetes Vorkommnis in der Dezimalbruchentwicklung von π, stützen können. Ich habe das vermieden, weil die Voraussetzung der Unentscheidbarkeit eines solchen Existenzproblems nicht nur völlig unberechtigt, sondern auch für den Beweis zu einem gewissen Grade unwesentlich ist. Wesentlich ist nur die Voraussetzung einer Unentscheidbarkeit überhaupt, eines „Ignorabimus“ in bezug auf Existenzprobleme der genannten Art.

(Eingegangen am 14. 10. 1929.)

Idealtheoretische Deutung der Darstellbarkeit beliebiger natürlicher Zahlen durch quadratische Formen[1]).

Von

Werner Weber in Göttingen.

Das Problem der Darstellbarkeit natürlicher Zahlen durch quadratische Formen mit ganzen rationalen Koeffizienten und von gegebener Klasse[2]) ist bisher nur auf dem Umweg über die Idealtheorie eines quadratischen Zahlkörpers mit Erfolg anzugreifen. Zum klassischen Bestande dieser Theorie gehört eine bei R. Dedekind[3]) und H. Weber[4]) vorkommende typische Relation, deren Bedeutung sich kurz durch das Schlagwort „darstellbar sein heißt Idealnorm sein" charakterisieren läßt[5]). Allen bisherigen Ergebnissen war nun gemeinsam, daß sie sich notwendigerweise auf den besonderen Fall beschränkten, in welchem die dargestellte Zahl bzw. die gegebene Idealnorm zu einem gewissen Bestandteil der Formendiskriminante, nämlich zu dem nach Abspaltung des Diskriminantenstammes Δ verbleibenden Faktor k^2 (k eine natürliche Zahl) teilerfremd ist. Im allgemeinen Fall versagt das Verfahren.

---

[1]) Die vorliegende Abhandlung ist bis auf geringfügige Änderungen ein Abdruck der von mir im April 1929 bei der mathematisch-naturwissenschaftlichen Fakultät der Universität Göttingen zur Erlangung der Doktorwürde eingereichten Dissertation. Ich habe Fräulein E. Noether für viele Ratschläge dabei zu danken.

[2]) Verzichtet man auf die Klassenbedingung, so wird das Problem arithmetisch wesentlich einfacher und ist als völlig gelöst zu betrachten. Man vergleiche etwa Dirichlet-Dedekind, Vorlesungen über Zahlentheorie, 4. Aufl. (1894), §§ 53—66 und § 91.

[3]) Es handelt sich um § 182 aus dem XI. Supplement des in Fußnote [2]) genannten Buches von Dirichlet und Dedekind. Das hier allein in Frage kommende XI. Supplement wird im folgenden unter „Dedekind" kurz mit „a. a. O." zitiert.

[4]) H. Weber, Lehrbuch der Algebra, 2. Aufl., **3** (1908), §§ 90—101. Im folgenden unter „H. Weber" mit „a. a. O." zitiert.

[5]) Man vergleiche auch Hecke, Theorie der algebraischen Zahlen (1923), Kap. VII, § 53; Fricke, Lehrbuch der Algebra **3** (1928), 2. Abschn., 3. Kap., § 6; Landau, Vorlesungen über Zahlentheorie **3** (1927), XI. Teil, Kap. 3, §§ 2, 3.

Zur Bewältigung dieses Falles kam aus der ganzen Literatur nur eine einzige, bei Dedekind[6]) auftretende Gleichung in Frage, die keine Teilerfremdheitsvoraussetzung enthält, aber andrerseits nicht von Idealen, sondern von beliebigen **Moduln** aus ganzen oder gebrochenen Zahlen des quadratischen Körpers handelt. Dedekind gewinnt für diese Moduln eine passende Normdefinition und erreicht damit eine scheinbar vollständige Analogie. Es zeigt sich nun freilich sofort, daß dieser Modulbegriff zu allgemein ist, um ein wirklich umkehrbar eindeutiges Verfahren zu liefern. Auch erscheint es recht unbequem, Moduln mit gebrochenen Elementen aufnehmen zu müssen; und überdies weiß man über die allgemeinsten Moduln, selbst aus ganzen Körperzahlen, noch recht wenig. **Der Zweck der nachstehenden Arbeit** ist es, **den Modulbegriff** hierbei **auf das geeignete Maß einzuschränken.** Es wird sich nämlich ergeben, daß man vollständig mit **Moduln aus ganzen Elementen**, und zwar speziell mit **Idealen** in den verschiedenen **Ordnungen** des quadratischen Körpers auskommt. Mehr noch: Die Ideale in einer Ordnung $\mathfrak{n}$ braucht man nicht sämtlich; sondern unter ihnen erscheinen diejenigen Ideale $\mathfrak{c}$ ausgezeichnet, die in keiner umfassenderen Ordnung mehr Ideal sind, d. h. die in einem von Dedekind präzisierten Sinne „der Gleichung $\mathfrak{c}^0 = \mathfrak{n}$ genügen" oder „die Ordnung $\mathfrak{n}$ haben". In der Hauptordnung ist diese einschränkende Bedingung natürlich von selbst erfüllt. Geht man aber in eine weniger umfassende Ordnung $\mathfrak{n}$ hinein, so spaltet sich diese Kategorie von Moduln aus ganzen Elementen sofort in zwei verschiedene auf: Moduln $\mathfrak{m}$ mit $\mathfrak{m}^0 = \mathfrak{n}$, die aber nicht notwendig Untermengen von $\mathfrak{n}$ sind, und Ideale $\mathfrak{m}$ in $\mathfrak{n}$, deren Ordnung $\mathfrak{m}^0$ eine echte oder unechte Obermenge von $\mathfrak{n}$ ist[7]). Der Durchschnitt dieser beiden Kategorien erscheint bereits algebraisch ausgezeichnet und ist bisher meines Wissens in der Literatur noch nirgends berücksichtigt worden.

Unter Benutzung dieses Begriffes wird sich zeigen, daß für die Darstellungen von Zahlen durch Formen der Diskriminante $k^2 \Delta$ lediglich die Ideale in derjenigen Ordnung $\mathfrak{n}$ des quadratischen Zahlkörpers von der Diskriminante Δ nötig sind, die den Führer k hat, unter diesen Idealen aber, wie gesagt, nur diejenigen, die zugleich von der Ordnung $\mathfrak{n}$ sind. Beschränkt man sich auf Formen einer gegebenen Klasse K, so gehört dazu eine gewisse Modulklasse M, in der die zugeordneten Ideale zu suchen sind. Wird endlich auch noch die darzustellende natürliche Zahl m vorgeschrieben, so kommen nur Ideale von der Norm m in Betracht. Jedem der so ausge-

---

[6]) A. a. O., § 187, (14).

[7]) Im „teilerfremden" Falle vereinfacht sich der Zusammenhang insofern, als dann die zweite Kategorie in der ersten steckt. Vgl. Satz 24.

siebten Ideale entsprechen nun, wie sich weiter zeigt, wirklich Darstellungen von m, und zwar vermittels eines genau angebbaren und rechnerisch brauchbaren Algorithmus. Dabei erscheinen bei gegebenem Ideal alle Formen aus K für die Darstellung gleichberechtigt. Aber auch innerhalb einer einzelnen Form können im allgemeinen dem gegebenen Ideal mehrere Darstellungen entsprechen. Es ist eine weitere Aufgabe, die in § 4 behandelt werden wird, die zu demselben Ideal gehörigen Darstellungen, sei es durch dieselbe oder durch verschiedene Formen, zusammenzufassen und ihre Verwandtschaft zahlentheoretisch zu fixieren. Bei festgehaltener Form werden sich diese Darstellungen ein-eindeutig den in der Ordnung $\mathfrak{n}$ gelegenen Einheiten von positiver Norm zuordnen. — Verallgemeinerungen sind dann dadurch möglich, daß man auch das Ideal variiert. Hierfür gibt § 6 ein naheliegendes Beispiel: Gefragt wird nach der Gesamtheit der Darstellungen, deren zugeordnete Ideale bis auf Einheitsfaktoren mit einem gegebenen Ideal übereinstimmen. Die Anzahl dieser Darstellungen läßt sich abermals mit Hilfe von Einheiten kennzeichnen, jedoch nicht in so einfacher Weise wie im bisherigen Fall.

Nach dieser Deutung der Darstellbarkeit ist es dann eine Aufgabe besonderer Art, die für eine allgemeine Ordnung gewonnenen Ergebnisse auf die Hauptordnung zu übertragen. Dieses Problem fällt aus dem Rahmen der vorliegenden Arbeit heraus; doch sollen hierüber in § 7 wenigstens einige erste Sätze bewiesen werden. Es wird sich dabei eine ausgezeichnete Stellung der von Grell[8]) so genannten „Erweiterungsideale" ergeben. Zugleich lassen sich unter Benutzung dieses Begriffes die klassischen Sätze über die Darstellung zu k teilerfremder Zahlen ohne Mühe einordnen.

§ 1.

Vorbemerkungen.

Eine primitive binäre quadratische Form mit ganzen rationalen Koeffizienten a, b, c, positivem Anfangskoeffizienten a und nichtquadratischer Diskriminante wird im folgenden kurz quadratische Form, noch kürzer Form genannt und, wenn es auf die Benennung der Variablen nicht ankommt, mit $\{a, b, c\}$ bezeichnet. Die Begriffe der Darstellbarkeit bzw. eigentlichen Darstellbarkeit einer Zahl durch eine Form sind bekannt, ebenso die einfachsten Eigenschaften der Formenklassen. Der letztere Begriff soll wiederum in eingeschränkter Bedeutung gebraucht werden: Eine Formenklasse bezeichnet im folgenden bei positiver Diskriminante jede beliebige, bei negativer Diskriminante dagegen nur jede positiv-definite Klasse. Geht

[8]) H. Grell, Beziehungen zwischen den Idealen verschiedener Ringe. Math. Annalen 97 (1927), S. 490–523.

die Form $ax^2+bxy+cy^2$ durch die ganzzahlige lineare Variablentransformation von der Determinante 1

$$\left\{\begin{aligned} x &= \alpha x' + \beta y', \\ y &= \gamma x' + \delta y' \end{aligned}\right. \tag{1}$$

in die Form $a'x'^2+b'x'y'+c'y'^2$ über, so soll von zwei zugehörigen Darstellungen

$$m = ax^2+bxy+cy^2,$$
$$m = a'x'^2+b'x'y'+c'y'^2$$

gesagt werden, daß sie durch die vollständige lineare Substitution (1) auseinander hervorgehen. Eine vollständige lineare Substitution einer Darstellung bezieht sich also auf Variablenwerte und Formen und ändert die dargestellte Zahl nicht.

Ist Δ eine von 1 verschiedene Stammdiskriminante[9]), so bedeutet künftig der „Körper Δ" den quadratischen Zahlkörper mit der Diskriminante Δ, hierin ein Modul eine additive Gruppe beliebiger ganzer oder gebrochener Körperzahlen. Ein Modul heißt ganz, wenn er nur ganze Zahlen enthält. Eine Modulbasis wird durch eckige Klammern ausgedrückt. Ist $\mathfrak{m}$ ein Modul mit endlicher Basis, so sei unter der Ordnung $\mathfrak{m}^0$ von $\mathfrak{m}$ der Ring derjenigen (eo ipso ganzen) Zahlen α des Körpers verstanden, für die der Modul $\alpha\mathfrak{m}$ eine Untermenge von $\mathfrak{m}$ ist. Die einfachsten Sätze über Ordnungen im quadratischen Körper können bei Dedekind[10]) nachgelesen werden; hier finde nur die Tatsache Platz, daß jede Ordnung $\mathfrak{n}$ eine Basisdarstellung von der Gestalt

$$\mathfrak{n} = [1, k\Theta]$$

zuläßt, worin Θ die ganze Zahl $\frac{\Delta+\sqrt{\Delta}}{2}$ und k eine durch $\mathfrak{n}$ eindeutig bestimmte natürliche Zahl, der „Führer" der Ordnung $\mathfrak{n}$, ist.

Der Modul aller Zahlen α, für die $\alpha\mathfrak{m}$ Untermenge von $\mathfrak{m}^0$ ist, wird mit $\mathfrak{m}^{-1}$ bezeichnet.

Für die späteren Untersuchungen ist der Begriff der Norm $N(\mathfrak{m})$ eines Moduls $\mathfrak{m}$ von zweigliedriger Basis besonders wichtig. In Übereinstimmung mit Dedekind[11]) soll hierunter der absolute Betrag der Determinante einer Linearsubstitution mit rationalen Koeffizienten verstanden werden, die eine Basis von $\mathfrak{m}^0$ in eine Basis von $\mathfrak{m}$ überführt.

---

[9]) Eine Diskriminante heißt Stammdiskriminante, wenn sie keinen quadratischen Teiler außer 1 hat, nach dessen Abspaltung eine Diskriminante übrigbleibt. Jede Diskriminante läßt sich eindeutig als Produkt einer Quadratzahl mit einer Stammdiskriminante darstellen.

[10]) A. a. O., S. 642.

[11]) A. a. O., S. 643.

Dedekind zeigt[12]), daß sich jede unverkürzbare zweigliedrige Modulbasis auf die Gestalt

$$[m, m\omega]$$

bringen läßt, worin m eine positive rationale Zahl ist. Ist $\mathfrak{m}$ der zugehörige Modul und bedeutet

$$ax^2 - bx + c = 0$$

die irreduzible quadratische Gleichung mit teilerfremden ganzen rationalen Koeffizienten und positivem Koeffizienten des höchsten Gliedes, der die Zahl ω genügt, so hat die Norm von $\mathfrak{m}$ den Wert

$$N(\mathfrak{m}) = \frac{m^2}{a}. \tag{2}$$

Zugleich ist

$$\mathfrak{m}^0 = [1, a\omega].$$

Bedeutet $\overline{\mathfrak{m}}$ denjenigen Modul, dessen Elemente konjugiert zu denen von $\mathfrak{m}$ sind, so gilt

$$\mathfrak{m}\overline{\mathfrak{m}} = N(\mathfrak{m})\,\mathfrak{m}^0. \tag{3}$$

Für zwei Moduln $\mathfrak{a}, \mathfrak{b}$ mit zweigliedriger Basis gilt immer die Gleichung

$$N(\mathfrak{a})\,N(\mathfrak{b}) = N(\mathfrak{a}\mathfrak{b}).$$

Die Einteilung der Moduln in Klassen und deren Zuordnung zu den Formenklassen wird im folgenden von Dedekind (a. a. O., S. 655 f.) übernommen und besteht darin, daß man der Form $\{a, b, c\}$ von der Diskriminante $D = k^2 \Delta$ die Klasse des Moduls

$$\left[1, \frac{b+\sqrt{D}}{2a}\right] \text{ [13])},$$

d. h. die Gesamtheit der von diesem Modul nur durch Zahlfaktoren positiver Norm unterschiedenen Moduln zuordnet. Jeder Modul dieser Klasse hat die Ordnung $\left[1, \frac{b+\sqrt{D}}{2}\right]$ vom Führer k.

Endlich wird alsbald auch der Begriff des Ideals in einer Ordnung herangezogen werden, der als hinlänglich bekannt gelten kann[14]).

[12]) Für das Folgende vgl. Dedekind, a. a. O., S. 640–645.

[13]) $\sqrt{D}$ soll hier und im folgenden, wie üblich, den positiven bzw. positiv-imaginären Wert der Quadratwurzel bedeuten.

[14]) Die von Dedekind in der Arbeit „Über die Anzahl der Idealklassen in den verschiedenen Ordnungen eines endlichen Körpers" (Festschrift zur Säkularfeier des Geburtstages von Carl Friedrich Gauß, 1877), § 4 gegebene Definition umfaßt nach heutigen Begriffen nur einen Spezialfall. Vgl. Fußnote [22]). — Wohl aber findet sich der Idealbegriff in dem auch heute üblichen Sinne in der 3. Auflage des Buches von Dirichlet und Dedekind (1879), § 172, S. 522. (In der 4. Auflage ist dieser Paragraph fortgefallen, und nur die Fußnote *) auf S. 554 deutet noch auf den allgemeineren Idealbegriff hin.)

An Bezeichnungen ist noch zu merken: Ist α eine Zahl des Körpers, so bedeutet $\bar{\alpha}$ die zu α konjugierte Zahl, $N(\alpha)$ die Norm von α. Derjenige Modul, dessen Elemente konjugiert zu denen des Moduls $\mathfrak{m}$ sind, heiße $\overline{\mathfrak{m}}$. Der Buchstabe $\mathfrak{o}$ bezeichnet stets die Hauptordnung. Mit (a, b) wird der größte gemeinsame Teiler der ganzen rationalen Zahlen a und b bezeichnet; a/b bedeutet „a ist Teiler von b"; $\subseteqq$ bedeutet Untermenge, $\subset$ echte Untermenge, $\supseteqq$ Obermenge, $\supset$ echte Obermenge, $\in$ das Enthaltensein, $\mathfrak{a} \frown \mathfrak{b}$ den Durchschnitt der Moduln $\mathfrak{a}$ und $\mathfrak{b}$.

§ 2.

Der $\mathfrak{m}$-Algorithmus.

Satz 1. *Die natürliche Zahl m ist dann und nur dann durch die Formenklasse K darstellbar, wenn es in der zugeordneten Idealklasse von der Ordnung* $\mathfrak{n}$ *ein Ideal in* $\mathfrak{n}$ *mit der Norm m gibt.*

Beweis. Es sei $\{a, b, c\}$ eine Form aus der Klasse K von der Diskriminante $D = k^2 \Delta$; x und y seien ganze rationale Zahlen und

$$m = ax^2 + bxy + cy^2. \tag{4}$$

Unter den Wurzeln

$$\frac{1}{a}\,\frac{b \pm \sqrt{D}}{2}$$

der Gleichung $au^2 - bu + c = 0$ sei diejenige, in welcher die Quadratwurzel das positive Vorzeichen hat, mit ω bezeichnet. In der K zugeordneten Modulklasse M, deren Ordnung $\mathfrak{n}$ den Führer k hat, liegt dann nach Definition der Modul $[1, \omega]$ und demnach auch der Modul

$$\mathfrak{c} = a(x + y\bar{\omega})[1, \omega];$$

denn die Norm des Zahlfaktors $a(x + y\bar{\omega})$ hat den Wert

$$a(x + y\bar{\omega})\,a(x + y\omega) = a(ax^2 + bxy + cy^2) = am,$$

ist also positiv. Da $a\omega$, $a\bar{\omega}$ und $a\omega\bar{\omega}$ in $\mathfrak{n}$ liegen, so ist $\mathfrak{c}$ Untermenge von $\mathfrak{n}$ (also insbesondere ein ganzer Modul). Weil aber $\mathfrak{c}$ die Ordnung $\mathfrak{n}$ hat, ist es sogar Ideal in $\mathfrak{n}$. Zugleich ist

$$N(\mathfrak{c}) = a\,m\,N([1, \omega]),$$

nach (2) hierin

$$N([1, \omega]) = \frac{1}{a},$$

also

$$N(\mathfrak{c}) = m \text{ [15]}.$$

[15]) Meine ursprüngliche Methode, zu diesem Ideal $\mathfrak{c}$ zu gelangen, war etwas umständlicher. Die hier gegebene Fassung verdanke ich einer Bemerkung von Fräulein E. Noether. Ähnliches gilt für den folgenden Rückweg von $\mathfrak{c}$ zur Darstellung.

Umgekehrt sei $\mathfrak{x}$ ein Ideal in $\mathfrak{n}$ von der Norm m, das zur Modulklasse M gehört. Wird zu irgendeiner Form $\{a, b, c\}$ aus K in der obigen Weise der in M gelegene Modul $[1, \omega]$ konstruiert, so liegt im Körper eine Zahl μ von positiver Norm, für die

$$\mu \mathfrak{x} = m[1, \omega]$$

ist. Nach (3) ist

$$\mu m \mathfrak{n} = \mu \mathfrak{x} \bar{\mathfrak{x}};$$

da mit $\mathfrak{x}$ auch $\bar{\mathfrak{x}}$ Untermenge von $\mathfrak{n}$ ist und $\mu \mathfrak{x}$ die Ordnung $\mathfrak{n}$ hat, so kommt

$$\mu m \mathfrak{n} \subseteq \mu \mathfrak{x} = m[1, \omega],$$

wegen $1 \in \mathfrak{n}$ also

$$\mu \in [1, \omega].$$

Es gibt demnach zwei ganze rationale Zahlen x und y mit

$$\mu = x + y\omega.$$

Wegen $N(\mu) > 0$ ist nun

$$N(\mu) N(\mathfrak{x}) = N(m) N([1, \omega]),$$

$$N(\mu) = \frac{m^2}{ma} = \frac{m}{a},$$

$$m\bar{\mu}[1, \omega] = \mu\bar{\mu}\mathfrak{x} = N(\mu)\mathfrak{x} = \frac{m}{a}\mathfrak{x},$$

$$\mathfrak{x} = a\bar{\mu}[1, \omega],$$

$$m = N(\mathfrak{x}) = a^2 \bar{\mu}\mu N([1, \omega]) = a(x + y\bar{\omega})(x + y\omega) = ax^2 + bxy + cy^2.$$

Das im ersten Teil des Beweises zur Konstruktion des Ideals $\mathfrak{c}$ in $\mathfrak{n}$ angewandte Verfahren ist offenbar bei gegebener Darstellung (4) völlig eindeutig. Es wird auch im weiteren Verlaufe dieser Arbeit eine Rolle spielen und verdient deshalb einen besonderen Namen. Der Modul $[1, \omega]$ wird künftig mit $\mathfrak{m}$, die Zahl $x + y\omega$ mit μ bezeichnet. Die zentrale Stellung des Moduls rechtfertigt die

Definition 1. *Die obige Konstruktion der Zahl ω, des Moduls $\mathfrak{m}$, der Zahl μ und des Ideals $\mathfrak{c}$ sei als der auf die Darstellung* (4) *angewandte $\mathfrak{m}$-Algorithmus bezeichnet. — Das Ideal $\mathfrak{c}$ heißt kurz das der Darstellung* (4) *zugeordnete Ideal.*

Die tiefere Bedeutung[16]) des Ideals $\mathfrak{c}$ erkennt man, wenn man $\mathfrak{c}$ in der allgemeinen Gestalt

$$\mathfrak{c} = (\varrho x + \sigma y)[1, \omega]$$

ansetzt, wo x und y die Variablenwerte der Ausgangsdarstellung bedeuten und ϱ und σ noch freibleibende Körperzahlen sind. Verlangt man, daß

[16]) Diese Überlegung rührt inhaltlich von Fräulein E. Noether her.

die Norm von $\mathfrak{c}$ in dieser Bezeichnungsweise formal gleich der gegebenen quadratischen Form in x und y wird, so ergibt sich:

$$\frac{1}{a}(\varrho x + \sigma y)(\bar{\varrho} x + \bar{\sigma} y) = ax^2 + bxy + cy^2$$
$$= a(x + \omega y)(x + \bar{\omega} y),$$

also bei passender Einheit ε

$$\varrho x + \sigma y = \varepsilon a(x + \omega y) \quad \text{oder} \quad \varrho x + \sigma y = \varepsilon a(x + \bar{\omega} y).$$

Damit $\mathfrak{c}$ Untermenge von $\mathfrak{n}$ ist, genügt es, den zweiten dieser Fälle zu betrachten und $\varepsilon = 1$ zu setzen, womit $\mathfrak{c}$ in der Tat in das oben konstruierte Ideal übergeht.

Der vorige Beweis liefert — da $\{a, b, c\}$ willkürlich in der Klasse K angenommen war — noch den

Satz 2. *Ist $\mathfrak{c}$ ein Ideal in $\mathfrak{n}$, das in der Klasse M liegt und die Norm m hat, so gibt es in jeder Form aus der M zugeordneten Formenklasse mindestens eine Darstellung von m, deren zugeordnetes Ideal $\mathfrak{c}$ ist.*

§ 3.
Der Teiler einer Darstellung.

In Analogie zu einem für den teilerfremden Fall gültigen Satz von H. Weber (a. a. O., § 97, Satz 8) kann man auch im allgemeinen Falle fragen, inwieweit sich die im größten gemeinsamen Teiler der darstellenden Variablenwerte aufgehenden Primfaktoren in den natürlichen Zahlteilern des zugeordneten Ideals wiederfinden und umgekehrt. Hierbei wird sich im folgenden eine ausgezeichnete Stellung der Primteiler des Führers ergeben.

Satz 3. *Es sei t eine von Null verschiedene ganze rationale Zahl. Dann haben die darstellenden Variablenwerte x und y in*

$$m = ax^2 + bxy + cy^2 \tag{4}$$

dann und nur dann den gemeinsamen Teiler t, wenn das zugeordnete Ideal $\mathfrak{c}$ durch t teilbar und überdies der Modul $\frac{\mathfrak{c}}{t}$ wieder Ideal in $\mathfrak{n}$ ist. Zugleich ist dann der Darstellung

$$\frac{m}{t^2} = a\left(\frac{x}{t}\right)^2 + b\frac{x}{t}\frac{y}{t} + c\left(\frac{y}{t}\right)^2 \tag{5}$$

das Ideal $\frac{\mathfrak{c}}{t}$ zugeordnet.

Beweis. 1. Es sei t/x, t/y. Die Buchstaben ω, $\mathfrak{m}$, μ mögen die Bedeutung aus dem auf (4) angewandten $\mathfrak{m}$-Algorithmus haben. Wendet man diesen auf die Darstellung (5) an, so bleiben ω und $\mathfrak{m}$ erhalten; μ wird ersetzt durch

$$\mu' = \frac{x}{t} + \frac{y}{t}\omega = \frac{\mu}{t}$$

und daher $\mathfrak{c}$ durch

$$\mathfrak{c}' = a\overline{\mu}'\mathfrak{m} = \frac{1}{t}\,a\overline{\mu}\mathfrak{m} = \frac{1}{t}\,\mathfrak{c},$$

so daß $\frac{\mathfrak{c}}{t}$ wieder Ideal in $\mathfrak{n}$ ist.

2. Umgekehrt sei $\mathfrak{c}' = \frac{\mathfrak{c}}{t}$ Ideal in $\mathfrak{n}$. Die Zwischenglieder des auf die Darstellung (4) angewandten $\mathfrak{m}$-Algorithmus seien ω, $\mathfrak{m}$, μ. Dann ist

$$\mathfrak{c}'\overline{\mathfrak{m}} \subseteqq \overline{\mathfrak{m}}, \tag{6}$$

da $\overline{\mathfrak{m}}$ die Ordnung $\mathfrak{n}$ hat und $\mathfrak{c}'$ in $\mathfrak{n}$ liegt. Die linke Seite hat aber wegen (3) den Wert

$$\left(\frac{x}{t} + \frac{y}{t}\,\overline{\omega}\right) a\,\mathfrak{m}\overline{\mathfrak{m}} = \left(\frac{x}{t} + \frac{y}{t}\,\overline{\omega}\right) a\,N(\mathfrak{m})\,\mathfrak{n}$$
$$= \left(\frac{x}{t} + \frac{y}{t}\,\overline{\omega}\right) \mathfrak{n}.$$

Da $\mathfrak{n}$ die Zahl 1 enthält, folgt also mit Rücksicht auf (6)

$$\frac{x}{t} + \frac{y}{t}\,\overline{\omega} \subseteqq \overline{\mathfrak{m}} = [1, \overline{\omega}],$$

so daß $\frac{x}{t}$ und $\frac{y}{t}$ ganz sein müssen.

Definition 2. *Das Ideal* $\mathfrak{c}$ *in der Ordnung* $\mathfrak{n}$ *habe einen ganzen rationalen Zahlteiler* t *von der Beschaffenheit, daß der Modul* $\frac{\mathfrak{c}}{t}$ *noch Ideal in* $\mathfrak{n}$ *ist. Dann heißt* t *ein (in bezug auf die Ordnung* $\mathfrak{n}$*) hebbarer ganzer rationaler Zahlteiler von* $\mathfrak{c}$.

Definition 3. *Der größte gemeinsame Teiler der Variablenwerte* x, y *in einer Darstellung*

$$m = ax^2 + bxy + cy^2$$

heiße der Teiler dieser Darstellung.

Aus Satz 3 ergibt sich dann sofort der

Satz 4. *Unter den hebbaren ganzen rationalen Zahlteilern jedes Moduls* $\mathfrak{c}$*, der Ideal in seiner Ordnung* $\mathfrak{c}^0$ *ist, gibt es* (natürlich) *einen größten; in diesem gehen alle übrigen auf. Er ist gleich dem Teiler jeder Darstellung, deren zugeordnetes Ideal* $\mathfrak{c}$ *ist.*

Satz 5. $\mathfrak{c}$ *sei Ideal in seiner Ordnung* $\mathfrak{c}^0 = \mathfrak{n}$ *mit dem Führer* k. *Dann ist jeder zu* k *teilerfremde ganze rationale Zahlteiler von* $\mathfrak{c}$ *hebbar (in bezug auf* $\mathfrak{n}$*).*

Beweis. Die ganze rationale Zahl t sei zu k teilerfremd und gehe in $\mathfrak{c}$ auf. Da der Modul $\frac{\mathfrak{c}}{t}$ die Ordnung $\mathfrak{n}$ hat, so genügt es, zu zeigen, daß er Untermenge von $\mathfrak{n}$ ist. Da $\mathfrak{c}$ diese Eigenschaft hat, so gibt es zu

jeder Zahl α aus $\mathfrak{c}$ eine ganze rationale Zahl r so, daß

$$\alpha \equiv r \pmod{k}$$

ist. Wegen der Teilerfremdheit von k und t ist ferner die Kongruenz

$$s\,t \equiv 1 \pmod{k}$$

durch ein ganzes rationales s lösbar. Setzt man $\alpha = t\alpha'$, so ist α' ganz, ferner

$$\alpha' \equiv st\alpha' \equiv s\alpha \equiv sr \pmod{k};$$

die Zahl α' ist also mod k kongruent einer ganzen rationalen Zahl und gehört mithin zu $\mathfrak{n}$. Damit ist die Behauptung bewiesen.

Hieraus folgen ohne weiteres die Sätze:

Satz 6. *Die Menge der zu k teilerfremden gemeinsamen Teiler der Variablenwerte einer Darstellung durch eine Form von der Diskriminante $k^2\Delta$ ist identisch mit der Menge der zu k teilerfremden ganzen rationalen Zahlteiler des zugeordneten Ideals in* $\mathfrak{n}$.

Satz 7. *Der Modul $\mathfrak{c}$ sei Ideal in seiner Ordnung $\mathfrak{c}^0 = \mathfrak{n}$ vom Führer k. Unter den zu k teilerfremden ganzen rationalen Zahlteilern von $\mathfrak{c}$ gibt es dann* (natürlich) *eine größte Zahl; in dieser gehen alle übrigen auf. Sie ist gleich dem größten zu k teilerfremden gemeinsamen Teiler der Variablenwerte jeder Darstellung, deren zugeordnetes Ideal $\mathfrak{c}$ ist.*

Satz 8. *Ist der Teiler einer Darstellung durch eine Form von der Diskriminante $k^2\Delta$ Potenzteiler*[17]) *von k, so ist jeder ganze rationale Zahlteiler des zugeordneten Ideals Potenzteiler von k.*

§ 4.

Assoziierte Darstellungen derselben Zahl.

Anschließend an das Ergebnis von § 2 erhebt sich sofort die Frage nach einer ein-eindeutigen Zuordnung. Zunächst ist aber klar, daß hierbei nur dann eine Hoffnung auf Eindeutigkeit bestehen kann, wenn man sich auf eine bestimmte Form aus der Klasse K beschränkt. Denn dasselbe Ideal $\mathfrak{c}$ ist nach Satz 2 mindestens einer Darstellung von m durch jede Form aus K zugeordnet. Es wird sich jedoch bald zeigen, daß auch eine Darstellung von m durch eine gegebene Form keineswegs durch Angabe ihres $\mathfrak{c}$ bestimmt ist.

Definition 4. *Zwei Darstellungen einer und derselben natürlichen Zahl durch zwei beliebige* (eventuell verschiedene) *Formen einer und derselben Klasse heißen assoziiert, wenn ihnen dasselbe Ideal zugeordnet ist.*

---

[17]) Eine Zahl a heißt Potenzteiler einer Zahl b, wenn sie in einer hinreichend hohen Potenz von b aufgeht, d. h. wenn alle verschiedenen Primfaktoren von a auch in b enthalten sind.

Eine formentheoretische Deutung erlangt dieser Begriff durch den

Satz 9. *Zwei Darstellungen einer und derselben natürlichen Zahl m durch Formen der Klasse K sind dann und nur dann assoziiert, wenn sie durch eine vollständige lineare Substitution von der Determinante* 1 *auseinander hervorgehen.*

Beweis. Die Variablensubstitution

$$\text{(1)}\qquad \begin{cases} x = \alpha x' + \beta y', \\ y = \gamma x' + \delta y' \end{cases}$$

mit $\alpha\delta - \beta\gamma = 1$ möge die Form $ax^2 + bxy + cy^2$ in die Form $a'x'^2 + b'x'y' + c'y'^2$ überführen. Auf ein spezielles Wertepaar x, y mit

$$\text{(4)}\qquad m = ax^2 + bxy + cy^2$$

angewandt, ergibt die Substitution (1) ein Wertepaar x', y' mit

$$\text{(7)}\qquad m = a'x'^2 + b'x'y' + c'y'^2.$$

In (4) und (7) hat man die allgemeine Gestalt zweier Darstellungen, die durch eine vollständige unimodulare Substitution ineinander übergehen. Der $\mathfrak{m}$-Algorithmus erzeuge nun aus den Darstellungen (4) und (7) die Zahlen ω und ω', die Moduln $\mathfrak{m}$ und $\mathfrak{m}'$, die Zahlen μ und μ' und die Ideale $\mathfrak{c}$ und $\mathfrak{c}'$; es ist zu zeigen, daß $\mathfrak{c} = \mathfrak{c}'$ ist. Zwischen den Werten ω und ω' besteht die Beziehung

$$\omega' = \frac{\beta + \delta\omega}{\alpha + \gamma\omega}.$$

Sind also X und Y ganze rationale Zahlen und ist

$$X = \alpha X' + \beta Y',$$
$$Y = \gamma X' + \delta Y',$$

so folgt identisch

$$X' + Y'\omega' = X' + Y'\frac{\beta + \delta\omega}{\alpha + \gamma\omega},$$

$$\text{(8)}\quad (\alpha + \gamma\omega)(X' + Y'\omega') = (\alpha X' + \beta Y') + (\gamma X' + \delta Y')\omega = X + Y\omega.$$

Da nun X', Y' mit X, Y alle Paare ganzer rationaler Zahlen darstellt, so besagt (8) dasselbe wie die Modulgleichung

$$\text{(9)}\qquad (\alpha + \gamma\omega)\mathfrak{m}' = \mathfrak{m}.$$

Setzt man speziell $X = x$, $Y = y$, so ergibt sich aus (8) weiter

$$(\alpha + \gamma\omega)\mu' = \mu,$$

demnach durch Übergang zum Konjugierten

$$\text{(10)}\qquad (\alpha + \gamma\overline{\omega})\overline{\mu'} = \overline{\mu}.$$

Durch Multiplikation folgt aus (9) und (10)

$$\overline{\mu}\,\mathfrak{m} = (\alpha + \gamma\omega)(\alpha + \gamma\overline{\omega})\,\overline{\mu'}\,\mathfrak{m}' = \left(\alpha^2 + \frac{b}{a}\alpha\gamma + \frac{c}{a}\gamma^2\right)\overline{\mu'}\,\mathfrak{m}',$$

also (wegen $a' = a\alpha^2 + b\alpha\gamma + c\gamma^2$)

$$\overline{\mu}\,\mathfrak{m} = \frac{a'}{a}\,\overline{\mu'}\,\mathfrak{m}',$$

$$a\overline{\mu}\,\mathfrak{m} = a'\overline{\mu'}\,\mathfrak{m}',$$

$$\mathfrak{c} = \mathfrak{c}'.$$

Beim Beweis der Umkehrung ist das Merkwürdige, daß die Behauptung auf Grund des ersten Teils fast trivial wird, wenn sie erst einmal für Darstellungen durch eine und dieselbe Form gezeigt ist. Ist sie nämlich für diesen Fall nachgewiesen, so schließt man im allgemeinen Falle folgendermaßen: Ist

(4) $$m = ax^2 + bxy + cy^2,$$

(7) $$m = a'x'^2 + b'x'y' + c'y'^2$$

und wird die Form $aX^2 + bXY + cY^2$ durch die Substitution

$$\begin{cases} X = \alpha X' + \beta Y', \\ Y = \gamma X' + \delta Y' \end{cases} \qquad (\alpha\delta - \beta\gamma = 1)$$

in die Form $a'X'^2 + b'X'Y' + c'Y'^2$ übergeführt (eine solche Substitution gibt es immer, wenn die beiden Formen zur selben Klasse gehören), so sei etwa

$$x = \alpha x'' + \beta y'',$$

$$y = \gamma x'' + \delta y''$$

gesetzt. Es ist dann

(11) $$m = a'x''^2 + b'x''y'' + c'y''^2.$$

Ist nun den Darstellungen (4) und (7) dasselbe Ideal $\mathfrak{c}$ zugeordnet, so entspricht $\mathfrak{c}$ nach dem schon bewiesenen Teil des Satzes auch der Darstellung (11), da (11) aus (4) durch eine vollständige lineare Substitution hervorgeht. Nach dem für eine und dieselbe Form als bewiesen angenommenen Satze geht also die Darstellung (7) aus der Darstellung (11) durch eine vollständige lineare Substitution hervor. Ist deren Matrix etwa $\begin{pmatrix}\alpha_1 & \beta_1 \\ \gamma_1 & \delta_1\end{pmatrix}$, so führt die vollständige Substitution mit der Matrix $\begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}\begin{pmatrix}\alpha_1 & \beta_1 \\ \gamma_1 & \delta_1\end{pmatrix}$ die Darstellung (4) in die Darstellung (7) über. Die Produktmatrix ist aber natürlich wieder ganzzahlig und von der Determinante 1.

Der Beweis braucht also nur noch für den Fall erbracht zu werden, daß zwei Darstellungen durch dieselbe Form $\{a, b, c\}$:

$$m = a x_1^2 + b x_1 y_1 + c y_1^2, \tag{12}$$

$$m = a x_2^2 + b x_2 y_2 + c y_2^2, \tag{13}$$

dasselbe Ideal $\mathfrak{c}$ zugeordnet ist. Der $\mathfrak{m}$-Algorithmus läßt aus den beiden Darstellungen dieselbe Zahl ω und denselben Modul $\mathfrak{m}$, sodann zwei Zahlen μ_1 und μ_2 entspringen. Dann wird

$$\mu_1 = x_1 + y_1 \omega,$$

$$\mu_2 = x_2 + y_2 \omega,$$

$$\mathfrak{c} = a \overline{\mu_1} \mathfrak{m} = a \overline{\mu_2} \mathfrak{m}. \tag{14}$$

Multiplikation von (14) mit der Hauptordnung $\mathfrak{o}$ ergibt

$$a \overline{\mu_1} \mathfrak{m} \mathfrak{o} = a \overline{\mu_2} \mathfrak{m} \mathfrak{o},$$

$$a \mu_1 \overline{\mathfrak{m}} \mathfrak{o} = a \mu_2 \overline{\mathfrak{m}} \mathfrak{o}.$$

Da $a \overline{\mathfrak{m}} \mathfrak{o}$ Ideal in $\mathfrak{o}$ ist, so muß es also in $\mathfrak{o}$ eine Einheit ε geben, für die

$$\mu_1 = \varepsilon \mu_2$$

ist. Wegen $N(\mu_1) = N(\mu_2) \left(= \frac{m}{a}\right)$ muß

$$N(\varepsilon) = +1$$

sein. Es folgt

$$\bar{\varepsilon}[1, \omega] = \bar{\varepsilon} \mathfrak{m} = \frac{\bar{\varepsilon} \mathfrak{c}}{a \overline{\mu_1}} = \frac{\mathfrak{c}}{a \overline{\mu_2}} = \mathfrak{m} = [1, \omega],$$

und demnach gibt es vier ganze rationale Zahlen $\alpha, \beta, \gamma, \delta$ mit $\alpha\delta - \beta\gamma = \pm 1$ so, daß

$$\begin{cases} \bar{\varepsilon} = \alpha + \beta \omega, \\ \bar{\varepsilon} \omega = \gamma + \delta \omega \end{cases} \tag{15}$$

ist. Setzt man

$$\begin{cases} x = \alpha x' + \gamma y', \\ y = \beta x' + \delta y', \end{cases} \tag{16}$$

so wird identisch

$$\begin{aligned} a x^2 + b x y + c y^2 &= a (x + y \omega)(x + y \overline{\omega}) \\ &= a ((\alpha + \beta \omega) x' + (\gamma + \delta \omega) y') ((\alpha + \beta \overline{\omega}) x' + (\gamma + \delta \overline{\omega}) y') \\ &= a \bar{\varepsilon} (x' + y' \omega) \varepsilon (x' + y' \overline{\omega}) \\ &= a (x' + y' \omega)(x' + y' \overline{\omega}), \end{aligned}$$

$$a x^2 + b x y + c y^2 = a x'^2 + b x' y' + c y'^2. \tag{17}$$

Da die Substitution (16) also die Form $\{a, b, c\}$ in sich überführt, muß ihre Determinante $\alpha\delta - \beta\gamma$ nach einem bekannten Satze den Wert $+1$ haben; dies folgt übrigens auch aus

$$1 = N(\varepsilon) = \varepsilon\bar{\varepsilon} = (\alpha + \beta\bar{\omega})\frac{\gamma + \delta\omega}{\omega},$$

$$\omega = \alpha\gamma + \alpha\delta\omega + \beta\gamma\bar{\omega} + \beta\delta\omega\bar{\omega}$$

$$= \alpha\gamma + (\alpha\delta - \beta\gamma)\omega + \beta\gamma\frac{b}{a} + \beta\delta\frac{c}{a}.$$

Andrerseits liefert das Einsetzen der speziellen Werte x_1, y_1 für x', y' in (16) wegen

$$x_2 + y_2\omega = \mu_2 = \frac{1}{\varepsilon}\mu_1 = \bar{\varepsilon}\mu_1 = \bar{\varepsilon}(x_1 + y_1\omega)$$

$$= (\alpha + \beta\omega)x_1 + (\gamma + \delta\omega)y_1$$

$$= (\alpha x_1 + \gamma y_1) + (\beta x_1 + \delta y_1)\omega$$

die Werte x_2, y_2. Mit Rücksicht auf (17) ergibt sich also, daß die Darstellung (12) aus der Darstellung (13) durch die vollständige lineare Substitution (16) hervorgeht, die die Form $\{a, b, c\}$ in sich überführt. Damit ist der Satz bewiesen.

Nebenbei läßt die zweite Hälfte dieses Beweises den folgenden Satz vermuten, der sich auch sofort bestätigen läßt:

Satz 10. *Ist $\{a, b, c\}$ eine Form von der Diskriminante $k^2 \Delta$, so lassen sich die in $\mathfrak{n}$ enthaltenen Einheiten von positiver Norm ein-eindeutig denjenigen Linearsubstitutionen zuordnen, welche die Form $\{a, b, c\}$ in sich überführen.*

Beweis. Irgendeiner Darstellung (4) durch die Form $\{a, b, c\}$ möge der $\mathfrak{m}$-Algorithmus die Zahl ω und den Modul $\mathfrak{m} = [1, \omega]$ zuweisen. Für jede Einheit ε aus $\mathfrak{n}$ mit $N(\varepsilon) = 1$ ist, da $\mathfrak{m}$ die Ordnung $\mathfrak{n}$ hat und auch $\bar{\varepsilon}$ in $\mathfrak{n}$ liegt,

$$\bar{\varepsilon}\mathfrak{m} \leqq \mathfrak{m},$$

andrerseits

$$\frac{1}{\bar{\varepsilon}}\mathfrak{m} = \varepsilon\mathfrak{m} \leqq \mathfrak{m},$$

$$\mathfrak{m} \leqq \bar{\varepsilon}\mathfrak{m},$$

also

$$\bar{\varepsilon}\mathfrak{m} = \mathfrak{m}.$$

Wiederum bestehen also zwei Gleichungen von der Gestalt (15), so daß die zugehörige Substitution (16) die Form $\{a, b, c\}$ in sich überführt. Der Einheit ε ordne man diese Substitution zu. Natürlich entsprechen verschiedenen Einheiten verschiedene Substitutionen. Andrerseits läßt sich jede Substitution von der Gestalt (16), welche die Form in sich überführt, auf die genannte Art aus einer in $\mathfrak{n}$ gelegenen Einheit ε von der Norm 1

erzeugen. Unter Zugrundelegung zweier beliebiger Darstellungen von der Gestalt (12) und (13), wobei (12) durch die vollständige Substitution (16) aus (13) hervorgeht, ergibt sich nämlich genau wie im zweiten Teil des Beweises von Satz 9 und unter Benutzung dieses Satzes die Existenz einer Einheit ε von positiver Norm mit $\bar{\varepsilon}\mathfrak{m} = \mathfrak{m}$, die definiert ist durch

$$x_1 + y_1\omega = \varepsilon(x_2 + y_2\omega).$$

Wegen $\bar{\varepsilon} \in \mathfrak{m}^0 = \mathfrak{n}$ liegt ε in $\mathfrak{n}$. Aus

$$x_2 = \alpha x_1 + \gamma y_1,$$
$$y_2 = \beta x_1 + \delta y_1$$

folgt

$$\bar{\varepsilon}(x_1 + y_1\omega) = (\alpha x_1 + \gamma y_1) + (\beta x_1 + \delta y_1)\omega = x_1(\alpha + \beta\omega) + y_1(\gamma + \delta\omega);$$

wegen der unschwer zu verifizierenden Gleichung $\gamma + \delta\omega = \omega(\alpha + \beta\omega)$ ist also

$$\bar{\varepsilon}(x_1 + y_1\omega) = (x_1 + y_1\omega)(\alpha + \beta\omega),$$
$$\bar{\varepsilon} = \alpha + \beta\omega,$$
$$\bar{\varepsilon}\omega = \gamma + \delta\omega,$$

so daß die Gleichungen (15) erfüllt sind. Damit ist der Satz völlig bewiesen.

Aus Satz 9 und 10 folgt noch:

Satz 11. *Diejenigen Darstellungen der natürlichen Zahl m durch die feste Form* $\{a, b, c\}$ *aus der Klasse K, denen das feste Ideal* $\mathfrak{c}$ *in* $\mathfrak{n}$ *aus der K zugeordneten Modulklasse zugeordnet ist, dessen Norm m ist, entsprechen ein-eindeutig denjenigen linearen Substitutionen von der Determinante* 1, *welche die Form* $\{a, b, c\}$ *in sich überführen. Ihre „Anzahl“* (worunter gegebenenfalls auch die „Zahl“ ∞ zu verstehen ist) *hängt also nur von der Diskriminante der Form ab,* im übrigen weder von der Form selbst noch von ihrer Klasse noch von der darzustellenden Zahl noch von dem Ideal $\mathfrak{c}$. *Sie ist überdies gleich der „Anzahl“ der in* $\mathfrak{n}$ *gelegenen Einheiten von der Norm* $+1$.

Der Anschluß des Assoziiertheitsbegriffes an den gleichlautenden alten arithmetischen Begriff kann nun in folgender Weise hergestellt werden:

Satz 12. *Die ganzzahlige Substitution*

$$\begin{cases} X = \alpha X' + \beta Y', \\ Y = \gamma X' + \delta Y' \end{cases}$$

habe die Determinante 1 *und führe die Form* $aX^2 + bXY + cY^2$ *in die Form* $mX'^2 + rX'Y' + sY'^2$ *über, so daß also insbesondere*

$$m = a\alpha^2 + b\alpha\gamma + c\gamma^2 \tag{18}$$

ist. Aus der letzteren Darstellung von m mögen durch den $\mathfrak{m}$-Algorithmus der Modul $\mathfrak{m}$ und die Zahl μ entstehen, aus der Darstellung

$$(19) \qquad m = m \cdot 1^2 + r \cdot 1 \cdot 0 + s \cdot 0^2$$

durch denselben Algorithmus der Modul $\mathfrak{r}$. Dann ist

$$\mu \mathfrak{r} = \mathfrak{m}.$$

Beweis. Sind ω und Ω durch den auf (18) bzw. (19) angewandten $\mathfrak{m}$-Algorithmus definiert, so ist

$$\mathfrak{m} = [1, \omega],$$
$$\mu = \alpha + \gamma\omega,$$
$$\Omega = \frac{\beta + \delta\omega}{\alpha + \gamma\omega},$$

daher

$$\mathfrak{r} = [1, \Omega] = \frac{1}{\alpha + \gamma\omega}[\alpha + \gamma\omega,\ \beta + \delta\omega] = \frac{1}{\mu}[\alpha + \gamma\omega,\ \beta + \delta\omega].$$

Da eine zweigliedrige Modulbasis stets ohne Änderung des Moduls einer ganzzahligen Linearsubstitution von der Determinante 1 unterworfen werden kann, so wird

$$\mathfrak{r} = \frac{1}{\mu}[1, \omega],$$
$$\mu \mathfrak{r} = \mathfrak{m}.$$

Satz 13. *Zwei eigentliche Darstellungen einer und derselben Zahl m durch eine und dieselbe Form $\{a, b, c\}$:*

$$(20) \qquad \begin{cases} m = a x_1^2 + b x_1 y_1 + c y_1^2, \\ m = a x_2^2 + b x_2 y_2 + c y_2^2, \end{cases}$$

sind dann und nur dann assoziiert, wenn unter den unendlich vielen Paaren ganzzahliger Linearsubstitutionen von der Determinante 1 *und der Gestalt*

$$(21) \qquad \begin{cases} X = x_1 X' + \beta_1 Y', \\ Y = y_1 X' + \delta_1 Y', \end{cases}$$

$$(22) \qquad \begin{cases} X = x_2 X' + \beta_2 Y', \\ Y = y_2 X' + \delta_2 Y' \end{cases}$$ [18])

mindestens eins (und daher jedes) *die Eigenschaft hat, daß die durch diese Substitutionen aus der Form $aX^2 + bXY + cY^2$ hervorgehenden*

[18]) Daß es solche Substitutionen gibt, und zwar in unendlicher Anzahl, folgt aus der Teilerfremdheit der darstellenden Variablenwerte in (20).

Formen von der Gestalt $mX'^2+r_1X'Y'+s_1Y'^2$ *und* $mX'^2+r_2X'Y'+s_2Y'^2$ *einander „parallel" sind, d. h. der Bedingung*

$$r_1 \equiv r_2 \pmod{2m} \tag{23}$$

genügen.

Beweis. 1. Die Darstellungen (20) seien assoziiert. Zwei unimodulare Substitutionen von der Gestalt (21) bzw. (22) mögen die Form $aX^2+bXY+cY^2$ in die Formen $mX'^2+r_1X'Y'+s_1Y'^2$ und $mX'^2+r_2X'Y'+s_2Y'^2$ überführen. Der $\mathfrak{m}$-Algorithmus ordnet den Darstellungen (20) einen gemeinsamen Modul $\mathfrak{m}$ und zwei Zahlen μ_1 und μ_2 von der Norm $\frac{m}{a}$, den Darstellungen

$$\left\{\begin{array}{l} m = m\cdot 1^2 + r_1\cdot 1\cdot 0 + s_1\cdot 0^2, \\ m = m\cdot 1^2 + r_2\cdot 1\cdot 0 + s_2\cdot 0^2 \end{array}\right. \tag{24}$$

zwei Zahlen Ω_1 und Ω_2 und zwei Moduln $\mathfrak{r}_1$ und $\mathfrak{r}_2$ zu. Nach Voraussetzung ist

$$\left\{\begin{array}{c} a\overline{\mu}_1\,\mathfrak{m} = a\overline{\mu}_2\,\mathfrak{m}, \\ \overline{\mu}_1\,\mathfrak{m} = \overline{\mu}_2\,\mathfrak{m}, \\ \frac{m}{a\mu_1}\,\mathfrak{m} = \frac{m}{a\mu_2}\,\mathfrak{m}, \\ \frac{\mathfrak{m}}{\mu_1} = \frac{\mathfrak{m}}{\mu_2}, \end{array}\right. \tag{25}$$

nach Satz 12 also

$$\mathfrak{r}_1 = \mathfrak{r}_2,$$
$$[1, \Omega_1] = [1, \Omega_2],$$
$$\Omega_1 \in [1, \Omega_2].$$

Es gibt also zwei ganze rationale Zahlen p und q so, daß

$$\Omega_1 = p + q\,\Omega_2$$

ist. Dies besagt:

$$\frac{1}{m}\,\frac{r_1+\sqrt{D}}{2} = p + q\,\frac{1}{m}\,\frac{r_2+\sqrt{D}}{2},$$

also

$$r_1 = 2mp + qr_2, \qquad 1 = q,$$
$$r_1 - r_2 = 2mp,$$
$$r_1 \equiv r_2 \pmod{2m}.$$

2. Umgekehrt sei für die Darstellungen (20) bei geeigneter Wahl der Substitutionen (21) und (22) die Bedingung (23) erfüllt. Durch den $\mathfrak{m}$-Algorithmus mögen aus den Darstellungen (24) die Zahlen Ω_1 und Ω_2

sowie die Moduln $\mathfrak{r}_1$ und $\mathfrak{r}_2$ hervorgehen. Dann ist

$$\Omega_1 = \frac{1}{m}\,\frac{r_1 + \sqrt{D}}{2},$$

$$\Omega_2 = \frac{1}{m}\,\frac{r_2 + \sqrt{D}}{2}.$$

Wegen (23) ist also

$$\Omega_1 - \Omega_2 = \frac{r_1 - r_2}{2m}$$

eine ganze Zahl; demnach besteht die Modulgleichung

$$[1, \Omega_1] = [1, \Omega_2],$$

$$\mathfrak{r}_1 = \mathfrak{r}_2.$$

Unter Benutzung von Satz 12 folgen hieraus in genau umgekehrter Reihenfolge die Relationen (25) und somit die Gleichheit der den Darstellungen (20) zugeordneten Ideale. Die Darstellungen sind also assoziiert.

Legt man in Satz 9 im Spezialfall eigentlicher Darstellungen durch eine und dieselbe Form die alte, als gleichberechtigt erwiesene arithmetische Definition des Assoziiertseins zugrunde, so bedarf der Satz des obigen Beweises nicht, sondern ist auch zahlentheoretisch leicht einzusehen. Die im Sinne des Beweises zu Satz 13 konstruierten Darstellungen (24) entsprechen sich nämlich vermittelst der vollständigen Substitution $\begin{pmatrix} 1 & \frac{r_2 - r_1}{2m} \\ 0 & 1 \end{pmatrix}$ bzw. ihrer Inversen, entsprechen aber andrerseits den gegebenen Darstellungen (20) vermittelst der vollständigen Substitutionen (21) und (22).

Aus Satz 4 folgt noch der

Satz 14. *Assoziierte Darstellungen einer und derselben Zahl haben denselben Teiler.*

§ 5.

Assoziierte Darstellungen verschiedener Zahlen.

Alles Bisherige bezog sich auf Darstellungen einer und derselben Zahl. Wünschenswert erscheint es nun, auch gewisse Darstellungen verschiedener Zahlen unter Umständen zu einer engeren Klasse vereinigen zu können, die sich durch ein und dasselbe in geeigneter Weise zugeordnete Ideal kennzeichnen läßt. Das Ideal $\mathfrak{c}$ ist hierzu unbrauchbar, da es die dargestellte Zahl m als seine Norm eindeutig bestimmt. Dagegen hilft eine andere Normierung einen Schritt weiter. Es sei nämlich d der Teiler einer

Darstellung, $\mathfrak{c}$ das zugeordnete Ideal, und man bilde den Modul

$$\mathfrak{s} = \frac{\mathfrak{c}}{d},$$

der nach Satz 3 wieder Ideal in $\mathfrak{n}$ ist[19]). Es gilt nun der

Satz 15. *Zwei Darstellungen einer und derselben natürlichen Zahl m durch beliebige Formen derselben Diskriminante D sind dann und nur dann assoziiert, wenn ihnen dasselbe Ideal $\mathfrak{s}$ entspricht.* Oder: *Wenn D und m gegeben sind, so besagt Gleichheit der $\mathfrak{c}$ dasselbe wie Gleichheit der $\mathfrak{s}$.*

Beweis. Es seien d_1 und d_2 die Teiler zweier Darstellungen von m, ferner $\mathfrak{c}_1$ und $\mathfrak{c}_2$ die zugeordneten Ideale, $\mathfrak{s}_1 = \frac{\mathfrak{c}_1}{d_1}$, $\mathfrak{s}_2 = \frac{\mathfrak{c}_2}{d_2}$. Sind zunächst die Darstellungen als assoziiert vorausgesetzt, so folgt aus Definition 4 und Satz 14 sofort $\mathfrak{s}_1 = \mathfrak{s}_2$. Ist umgekehrt $\mathfrak{s}_1 = \mathfrak{s}_2$ vorausgesetzt, so ergibt sich hieraus durch Übergang zur Norm

$$\frac{m}{d_1^2} = \frac{m}{d_2^2},$$

$$d_1 = d_2,$$

$$\mathfrak{c}_1 = d_1 \mathfrak{s}_1 = d_2 \mathfrak{s}_2 = \mathfrak{c}_2;$$

die Darstellungen sind also assoziiert.

Dieser Satz rechtfertigt die folgende erweiterte

Definition 5. *Zwei Darstellungen beliebiger natürlicher Zahlen durch beliebige Formen derselben Diskriminante heißen* assoziiert, *wenn ihnen dasselbe Ideal $\mathfrak{s}$ entspricht.*

Der arithmetische Übergang zu eigentlichen Darstellungen kann nun folgendermaßen vollzogen werden:

Satz 16. *Haben in der Darstellung*

$$m = ax^2 + bxy + cy^2 \tag{4}$$

die Variablenwerte x und y den gemeinsamen Teiler t, so entspricht der Darstellung

$$\frac{m}{t^2} = a\left(\frac{x}{t}\right)^2 + b\frac{x}{t}\frac{y}{t} + c\left(\frac{y}{t}\right)^2 \tag{5}$$

dasselbe Ideal $\mathfrak{s}$ wie der Darstellung (4).

---

[19]) Ursprünglich arbeitete ich hier mit dem komplizierteren Modul $\frac{d}{N(\mathfrak{c})}\mathfrak{c}$, der im Falle einer eigentlichen Darstellung mit dem Modul $\mathfrak{r}$ aus Satz 12 übereinstimmt. Die Einführung des Ideals $\mathfrak{s}$ schlug mir Fräulein E. Noether vor.

Beweis. Der Darstellung (4) sei das Ideal $\mathfrak{c}$, der Darstellung (5) also nach Satz 3 das Ideal $\frac{\mathfrak{c}}{t}$ zugeordnet. Bedeutet d den Teiler von (4), so ist

$$\frac{\frac{\mathfrak{c}}{t}}{\frac{d}{t}} = \frac{\mathfrak{c}}{d},$$

und das ist die Behauptung.

Satz 17. *Zwei Darstellungen beliebiger natürlicher Zahlen:*

$$(26)\qquad \begin{cases} m_1 = a_1 x_1^2 + b_1 x_1 y_1 + c_1 y_1^2, \\ m_2 = a_2 x_2^2 + b_2 x_2 y_2 + c_2 y_2^2, \end{cases}$$

von den Teilern d_1 und d_2 sind dann und nur dann assoziiert, wenn die eigentlichen Darstellungen

$$(27)\qquad \begin{cases} \frac{m_1}{d_1^2} = a_1 x_1'^2 + b_1 x_1' y_1' + c_1 y_1'^2, \\[2mm] \frac{m_2}{d_2^2} = a_2 x_2'^2 + b_2 x_2' y_2' + c_2 y_2'^2 \end{cases}$$

mit

$$x_1' = \frac{x_1}{d_1}, \quad y_1' = \frac{y_1}{d_1},$$

$$x_2' = \frac{x_2}{d_2}, \quad y_2' = \frac{y_2}{d_2}$$

assoziiert sind. Zugleich ist dann $\frac{m_1}{d_1^2} = \frac{m_2}{d_2^2}$.

Beweis. Der erste Teil des Satzes folgt nach Definition 5 sofort aus Satz 16. Die Gleichung $\frac{m_1}{d_1^2} = \frac{m_2}{d_2^2}$ ergibt sich dann aus der Gleichheit der $\mathfrak{z}$ durch Übergang zur Norm.

§ 6.

Vereinigte Darstellungen.

Nächst den assoziierten Darstellungen derselben Zahl, deren zugehörige Ideale $\mathfrak{c}$ miteinander übereinstimmen, beanspruchen diejenigen Scharen von Darstellungen ein erhöhtes Interesse, deren zugehörige $\mathfrak{c}$ sich nur durch Einheitsfaktoren unterscheiden.

Definition 6. *Zwei Darstellungen einer und derselben natürlichen Zahl heißen* vereinigt, *wenn die zugeordneten Ideale in der Ordnung $\mathfrak{n}$ sich nur durch einen Einheitsfaktor unterscheiden.*

Es gilt also:

Satz 18. *Assoziierte Darstellungen einer und derselben natürlichen Zahl sind stets vereinigt.*

Satz 19. *Stehen zwei Ideale* $\mathfrak{c}_1$ *und* $\mathfrak{c}_2$ *in* $\mathfrak{n}$, *deren Ordnung genau* $\mathfrak{n}$ *ist*[20]), *zueinander in der Beziehung*

$$\mathfrak{c}_1 = \varepsilon\,\mathfrak{c}_2,$$

wo ε *eine Körpereinheit ist, so ist dann und nur dann* $\mathfrak{c}_1 = \mathfrak{c}_2$, *wenn* ε *in* $\mathfrak{n}$ *liegt.*

Beweis. 1. Es sei $\mathfrak{c}_1 = \mathfrak{c}_2$. Dann folgt $\mathfrak{c}_1 = \varepsilon\,\mathfrak{c}_1$, nach Definition der Ordnung eines Moduls also $\varepsilon \in \mathfrak{c}_1^0$, $\varepsilon \in \mathfrak{n}$.

2. Es sei $\varepsilon \in \mathfrak{n}$. Da $\mathfrak{c}_2$ Ideal in $\mathfrak{n}$ ist, so wird $\varepsilon\,\mathfrak{c}_2 \leqq \mathfrak{c}_2$, $\mathfrak{c}_1 \leqq \mathfrak{c}_2$. Da nun ε die Norm ± 1 hat, so ist

$$\bar{\varepsilon}\,\mathfrak{c}_1 = \bar{\varepsilon}\,\varepsilon\,\mathfrak{c}_2 = \mathfrak{c}_2. \tag{28}$$

Mit ε liegt aber auch $\bar{\varepsilon}$ in der Ordnung $\mathfrak{n}$. Aus (28) folgt also ebenso $\mathfrak{c}_2 \leqq \mathfrak{c}_1$, also $\mathfrak{c}_1 = \mathfrak{c}_2$.

Die hiernach zu vermutende ein-eindeutige Zuordnung der in $\mathfrak{n}$ enthaltenen Einheiten zu den Elementen einer Schar assoziierter Darstellungen einer festen natürlichen Zahl durch eine feste Form besteht nach Satz 11 nur dann, wenn es in $\mathfrak{n}$ keine Einheiten negativer Norm gibt.

Satz 20. *Ist* $\mathfrak{c}$ *Ideal in* $\mathfrak{n}$ *von der Ordnung* $\mathfrak{n}$, *ferner* ε *eine Einheit, so ist* $\varepsilon\,\mathfrak{c}$ *dann und nur dann Ideal in* $\mathfrak{n}$, *wenn* ε *in* $\mathfrak{c}^{-1}$ *liegt.*

Beweis. 1. $\varepsilon\,\mathfrak{c}$ sei Ideal in $\mathfrak{n}$. Dann ist $\varepsilon\,\mathfrak{c} \leqq \mathfrak{n} = \mathfrak{c}^0$, nach Definition von $\mathfrak{c}^{-1}$ (§ 1) also $\varepsilon \in \mathfrak{c}^{-1}$.

2. Die Einheit ε liege in $\mathfrak{c}^{-1}$. Dann ist $\varepsilon\,\mathfrak{c} \leqq \mathfrak{c}^0 = \mathfrak{n}$, und da $\varepsilon\,\mathfrak{c}$ die Ordnung $\mathfrak{n}$ hat, so ist es sogar Ideal in $\mathfrak{n}$.

Satz 21. *Ist* $\mathfrak{c}$ *Ideal in* $\mathfrak{n}$ *von der Ordnung* $\mathfrak{n}$, *so ist die Anzahl*[21]) *derjenigen Ideale in* $\mathfrak{n}$, *die sich von* $\mathfrak{c}$ *nur durch Einheitsfaktoren unterscheiden, gleich dem Index der multiplikativen Gruppe* $\mathfrak{E}'$ *der in* $\mathfrak{n}$ *enthaltenen Einheiten in der Obergruppe* $\mathfrak{E}$ *der in* $\mathfrak{c}^{-1}$ *enthaltenen Einheiten.*

(Man beachte, daß wegen $\mathfrak{c}\,\mathfrak{n} = \mathfrak{c} \leqq \mathfrak{n}$ stets $\mathfrak{c}^{-1} \geqq \mathfrak{n}$ ist.)

Beweis. Nach Satz 20 sind die in Rede stehenden Ideale genau die Moduln von der Gestalt $\varepsilon\,\mathfrak{c}$, wo ε eine Einheit aus $\mathfrak{c}^{-1}$, d. h. Element von $\mathfrak{E}$

---

[20]) Zwischen Idealen in der Ordnung $\mathfrak{n}$ und Idealen (oder Moduln) von der Ordnung $\mathfrak{n}$ muß scharf geschieden werden. Vgl. die Einleitung.

[21]) Darunter ist, wie bei Satz 11, gegebenenfalls auch die „Zahl“ ∞ zu verstehen.

ist. Sind ε_1 und ε_2 Elemente von $\mathfrak{E}$, so ist dann und nur dann $\varepsilon_1 \mathfrak{c} = \varepsilon_2 \mathfrak{c}$, wenn $\mathfrak{c} = \frac{\varepsilon_2}{\varepsilon_1} \mathfrak{c}$ ist, d. h. nach Satz 19: wenn $\frac{\varepsilon_2}{\varepsilon_1}$ in $\mathfrak{E}'$ liegt oder wenn ε_1 und ε_2 zur selben Restklasse von $\mathfrak{E}$ nach $\mathfrak{E}'$ gehören. Daraus folgt die Behauptung.

Aus der in § 5 gegebenen Definition des Ideals $\mathfrak{z}$ folgt nun sofort der

Satz 22. *Zwei Darstellungen einer und derselben natürlichen Zahl m vom selben Teiler d sind dann und nur dann vereinigt, wenn die ihnen zugeordneten Ideale $\mathfrak{z}$ sich nur durch einen Einheitsfaktor unterscheiden.* Oder: *Bei gegebenem D, m und d stimmen dann und nur dann die $\mathfrak{c}$ bis auf einen Einheitsfaktor überein, wenn die $\mathfrak{z}$ es tun.*

Dieser Satz rechtfertigt die gegenüber Definition 6 zum Teil erweiterte, zum Teil dahinter zurückbleibende

Definition 7. *Haben zwei Darstellungen natürlicher Zahlen m_1 und m_2 die Teiler d_1 bzw. d_2 und ist*

$$\frac{m_1}{d_1^2} = \frac{m_2}{d_2^2}, \tag{29}$$

so heißen die Darstellungen vereinigt, *wenn die zugehörigen Ideale $\mathfrak{z}$ sich nur durch einen Einheitsfaktor unterscheiden.*

In den Fällen nämlich, auf die sowohl Definition 6 wie auch Definition 7 paßt, ist $m_1 = m_2$, nach (29) also $d_1 = d_2$; Satz 22 lehrt also die Äquivalenz der beiden Definitionen.

Satz 23. *Assoziierte Darstellungen beliebiger natürlicher Zahlen m_1, m_2 sind stets vereinigt.*

Beweis. Sind d_1 und d_2 die Teiler der Darstellungen, so gilt nach Satz 17 die Gleichung (29); die Darstellungen fallen also unter Definition 7. Aus Definition 5 folgt also die Behauptung.

Da für den Schlußsatz dieses Paragraphen eine Hilfsbetrachtung über „Ideale in verschiedenen Ringen" (Satz 25) nötig ist und diese Methode in § 7 nochmals zur Geltung kommt, so ist es zweckmäßig, an dieser Stelle die beiden folgenden Hilfssätze einzufügen, deren erster nur in § 7 benutzt wird.

Satz 24. *Ist $\mathfrak{a}$ ein zu k teilerfremdes Ideal in $\mathfrak{o}$, ferner*

$$\mathfrak{x} = \mathfrak{a} \frown \mathfrak{n},$$

so ist $\mathfrak{x}$ Ideal in $\mathfrak{n}$, genügt der Gleichung

$$\mathfrak{x}\mathfrak{o} = \mathfrak{a}$$

und hat die Ordnung $\mathfrak{n}$.

Beweis. Hinsichtlich der beiden ersten Behauptungen kann auf § 5 der in Fußnote [14]) genannten Arbeit von Dedekind verwiesen werden [22]). Zu zeigen ist also nur noch $\mathfrak{x}^0 = \mathfrak{n}$. Es sei $\mathfrak{x}^0 = \mathfrak{n}_1$ gesetzt; da $\mathfrak{x}$ Ideal in $\mathfrak{n}$ ist, so ist $\mathfrak{n}_1 \geqq \mathfrak{n}$. Bezeichnet man die zu k teilerfremde Zahl $N(\mathfrak{x}) = N(\mathfrak{a})$ mit m, so ist nach (3)

$$m\mathfrak{n}_1 = \mathfrak{x}\bar{\mathfrak{x}};$$

da $\mathfrak{x}$ Ideal in $\mathfrak{n}$, $\bar{\mathfrak{x}}$ Untermenge von $\mathfrak{n}$ ist, so folgt

$$m\mathfrak{n}_1 \leqq \mathfrak{x},$$
$$m \in \mathfrak{x}.$$

Jedes Element α von $\mathfrak{n}_1$ genügt demnach der Bedingung

$$\alpha m \in \mathfrak{x} \leqq \mathfrak{n}.$$

In

$$\alpha = \frac{\alpha m}{m}$$

sind also Zähler und Nenner des Bruches Zahlen aus der Ordnung $\mathfrak{n}$, der Nenner überdies zu k teilerfremd; da nun α ganz ist, so muß es sogar in $\mathfrak{n}$ liegen [23]). Somit ist $\mathfrak{n}_1 \leqq \mathfrak{n}$, also $\mathfrak{n}_1 = \mathfrak{n}$.

Satz 25. *Jedes Ideal $\mathfrak{c}$ in $\mathfrak{n}$ von der Ordnung $\mathfrak{n}$, dessen Norm m zu k teilerfremd ist, genügt der Gleichung*

$$\mathfrak{c} = \mathfrak{c}\mathfrak{o} \wedge \mathfrak{n}.$$

Beweis. Nach (3) ist

$$m\mathfrak{n} = \mathfrak{c}\bar{\mathfrak{c}} \leqq \mathfrak{c};$$

wegen $(m, k) = 1$ ist

$$m\mathfrak{n} + k\mathfrak{o} \geqq m\mathfrak{n} + k\mathfrak{n} = \mathfrak{n},$$

also um so mehr

$$\mathfrak{c} + k\mathfrak{o} \geqq \mathfrak{n},$$
$$\mathfrak{c} + k\mathfrak{o} = \mathfrak{n},$$

und § 5 der mehrfach zitierten Arbeit von Dedekind liefert die Behauptung.

Satz 26. *Zwei Darstellungen natürlicher Zahlen m_1, m_2 durch Formen der Diskriminante $k^2\Delta$ mögen die Teiler d_1 bzw. d_2 haben. Es sei*

$$\frac{m_1}{d_1^2} = \frac{m_2}{d_2^2},$$

---

[22]) Dedekinds Begriff des „Ideals in der Ordnung $\mathfrak{n}$" ist dort freilich ein anderer als der heute gebräuchliche: Außer den gewöhnlichen Idealeigenschaften verlangt Dedekind von einem solchen Ideal $\mathfrak{x}$ noch das Bestehen der Gleichung $\mathfrak{x} + k\mathfrak{o} = \mathfrak{n}$. Aber jedes Ideal im dortigen Sinne ist auch Ideal in dem oben zugrunde gelegten Sinne, und das genügt für den vorliegenden Zweck. — Man findet das Ergebnis auch bei Grell, a. a. O., S. 516.

[23]) Vgl. den Beweis zu Satz 5.

und der gemeinsame Wert dieser beiden Brüche sei zu k teilerfremd. Sind dann die Darstellungen vereinigt, so sind sie sogar assoziiert.

Beweis. Zunächst sei $m_1 = m_2$, $d_1 = d_2 = 1$. Dann ist also $m_1 (= m_2)$ zu k teilerfremd. Sind $\mathfrak{c}_1$ und $\mathfrak{c}_2$ die den Darstellungen zugeordneten Ideale, so gibt es nach Definition 6 eine Einheit ε so, daß

$$\mathfrak{c}_1 = \varepsilon \mathfrak{c}_2$$

ist. Es folgt

$$\mathfrak{c}_1 \mathfrak{o} = \varepsilon \mathfrak{c}_2 \mathfrak{o} = \mathfrak{c}_2 \mathfrak{o},$$

also nach Satz 25

$$\mathfrak{c}_1 = \mathfrak{c}_1 \mathfrak{o} \wedge \mathfrak{n} = \mathfrak{c}_2 \mathfrak{o} \wedge \mathfrak{n} = \mathfrak{c}_2;$$

die Darstellungen sind also assoziiert.

Im allgemeinen Fall seien etwa die Darstellungen (26) vorgelegt. Entsprechen ihnen die Ideale $\mathfrak{z}_1$ und $\mathfrak{z}_2$, den zugehörigen Darstellungen (27) die Ideale $\mathfrak{z}_1'$ und $\mathfrak{z}_2'$, so ist nach Satz 16

$$\mathfrak{z}_1' = \mathfrak{z}_1,$$
$$\mathfrak{z}_2' = \mathfrak{z}_2;$$

nach Definition 7 sind also auch die Darstellungen (27) vereinigt, also, da sie dem schon erledigten Fall angehören, sogar assoziiert. Nach Satz 17 sind dann auch die Darstellungen (26) assoziiert.

Im Falle $\Delta < -4$ sind übrigens vereinigte Darstellungen stets assoziiert; das erkennt man direkt aus Definition 6 bzw. 7, da es in diesem Fall außer ± 1 keine Einheiten gibt.

§ 7.

Bedeutung der Erweiterungsideale in der Hauptordnung.

In der in Fußnote [8]) genannten Arbeit von Grell findet sich der Begriff des „Erweiterungsideals" präzisiert. Für den hier vorliegenden Fall des Systems der beiden Ringe $\mathfrak{n}$ und $\mathfrak{o}$ sei aus der genannten Arbeit die folgende Definition entnommen:

Definition 8. *Ein Ideal $\mathfrak{x}$ im Ring $\mathfrak{o}$ heißt* **Erweiterungsideal** *(in bezug auf $\mathfrak{n}$), wenn es sich in der Gestalt $\mathfrak{y}\mathfrak{o}$ darstellen läßt, wo $\mathfrak{y}$ Ideal in $\mathfrak{n}$ ist.*

Satz 27. *Ist die natürliche Zahl m durch die Formenklasse K darstellbar, M die K zugeordnete Modulklasse, O die Klasse des Moduls $\mathfrak{o}$, so gibt es in der Klasse MO ein Erweiterungsideal (in $\mathfrak{o}$) von der Norm m.*

Beweis. Zu einer Darstellung von m durch eine Form aus K werde im Sinne des $\mathfrak{m}$-Algorithmus das Ideal $\mathfrak{c}$ in M von der Norm m konstruiert. Das Ideal $\mathfrak{c}\mathfrak{o}$ in $\mathfrak{o}$ liegt dann in der Klasse MO, hat die Norm m und ist Erweiterungsideal.

Dieser Satz ist aber nicht allgemein umkehrbar. Aus der Existenz eines Erweiterungsideals $\mathfrak{x}$ von der Norm m in MO kann nämlich nicht rückwärts auf die Existenz eines Ideals $\mathfrak{y}$ in $\mathfrak{n}$ von der Norm m geschlossen werden, das in M liegt. Um ein solches Ideal zu konstruieren, ist man versucht, es noch der Bedingung $\mathfrak{y}\mathfrak{o} = \mathfrak{x}$ zu unterwerfen, die erfüllbar scheint, weil $\mathfrak{x}$ Erweiterungsideal ist. Die Forderung, daß $\mathfrak{y}$ speziell der Klasse M angehört, ist dabei jedoch zu stark[24]); ein Ideal $\mathfrak{y}$ in $\mathfrak{n}$ mit $\mathfrak{y}\mathfrak{o} = \mathfrak{x}$ braucht nicht einmal die Ordnung $\mathfrak{n}$ zu haben[25]).

Dieser Übelstand läßt sich nun aber in gewissem Grade reparieren. Im allgemeinen gibt es nämlich, wie Grell[26]) im einzelnen ausführt, zu einem gegebenen Erweiterungsideal $\mathfrak{x}$ mehrere Ideale $\mathfrak{y}$ in $\mathfrak{n}$ mit $\mathfrak{y}\mathfrak{o} = \mathfrak{x}$. Für ihre Gesamtheit werde von Grell die folgende Bezeichnung übernommen:

Definition 9. *Die Gesamtheit aller Ideale $\mathfrak{y}$ in $\mathfrak{n}$, für die $\mathfrak{y}\mathfrak{o}$ gleich dem festen Erweiterungsideal $\mathfrak{x}$ in $\mathfrak{o}$ ist, heißt die zu $\mathfrak{x}$ gehörige Erweiterungsklasse*[27]) (*in der Ordnung* $\mathfrak{n}$).

Offenbar muß die Ordnung jedes Ideals der Erweiterungsklasse die Ordnung $\mathfrak{n}$ umfassen. Es gilt nun aber der

Satz 28. *Ist $\mathfrak{y}$ Ideal in $\mathfrak{n}$ und von einer Ordnung $\mathfrak{n}_1 \supset \mathfrak{n}$, so gibt es in der zu $\mathfrak{y}\mathfrak{o}$ gehörigen Erweiterungsklasse ein Ideal $\mathfrak{z}$, welches echtes Vielfaches von $\mathfrak{y}$ ist und genau die Ordnung $\mathfrak{n}$ hat.*

Beweis. Wegen $\mathfrak{n}\mathfrak{n}_1 = \mathfrak{n}_1$ gibt es nach Dedekinds Theorie der Moduln[28]) einen Modul $\mathfrak{z}$ von der Ordnung $\mathfrak{n}$, der die Gleichung

$$\mathfrak{z}\mathfrak{n}_1 = \mathfrak{y}$$

löst. Für dieses $\mathfrak{z}$ wird zunächst

$$\mathfrak{z} \subseteq \mathfrak{z}\mathfrak{n}_1 = \mathfrak{y},$$

wegen $\mathfrak{z}^0 \neq \mathfrak{y}^0$ aber sogar

$$\mathfrak{z} \subset \mathfrak{y}.$$

Hieraus folgt überdies $\mathfrak{z} \subset \mathfrak{n}$; wegen $\mathfrak{z}^0 = \mathfrak{n}$ ist also $\mathfrak{z}$ Ideal in $\mathfrak{n}$. Endlich ist

$$\mathfrak{z}\mathfrak{o} = \mathfrak{z}\mathfrak{n}_1\mathfrak{o} = \mathfrak{y}\mathfrak{o};$$

$\mathfrak{z}$ liegt also in der zu $\mathfrak{y}\mathfrak{o}$ gehörigen Erweiterungsklasse aus $\mathfrak{n}$.

---

[24]) Sie ließe sich natürlich erfüllen, wenn $\mathfrak{y}$ nur Modul zu sein brauchte; aber $\mathfrak{y}$ soll Ideal in $\mathfrak{n}$ sein.

[25]) Ein einfaches Beispiel hierfür bietet schon der Führer $\mathfrak{y} = k\mathfrak{o}$. — Ist dagegen $N(\mathfrak{x})$ zu k teilerfremd angenommen, so muß allerdings, wie sich zeigen läßt, $\mathfrak{y}^0 = \mathfrak{n}$ sein.

[26]) A. a. O., S. 507.

[27]) Natürlich hat dieser Begriff nichts mit Idealklassen im gewöhnlichen Sinne zu tun.

[28]) A. a. O., S. 650ff.

Hierin steckt unter anderem der

Satz 29. *In jeder Erweiterungsklasse aus* $\mathfrak{n}$ *gibt es mindestens ein Ideal, dessen Ordnung genau* $\mathfrak{n}$ *ist.*

Man greife nämlich irgendein Ideal $\mathfrak{y}$ aus der Erweiterungsklasse heraus. Hat $\mathfrak{y}$ schon die Ordnung $\mathfrak{n}$, so ist man fertig; andernfalls liefert der vorige Satz die Behauptung.

Mit diesen Hilfsmitteln gelingt nun eine teilweise Umkehrung des Satzes 27 in folgender Weise:

Satz 30. *Gibt es in der Idealklasse A in der Hauptordnung ein Erweiterungsideal $\mathfrak{x}$ von der Norm m und sind $M_1, M_2, \ldots, M_r$ diejenigen Modulklassen von der Ordnung* $\mathfrak{n}$, *die durch Multiplikation mit der Klasse O des Moduls* $\mathfrak{o}$ *die Klasse AO liefern, $K_1, K_2, \ldots, K_r$ die zugehörigen Formenklassen, so ist m durch mindestens eine der Klassen K_i darstellbar.*

Beweis. Nach Satz 29 gibt es in der zu $\mathfrak{x}$ gehörigen Erweiterungsklasse in $\mathfrak{n}$ mindestens ein Ideal $\mathfrak{y}$, das genau die Ordnung $\mathfrak{n}$ hat, also in einer Modulklasse von der Ordnung $\mathfrak{n}$ liegt. Wegen $\mathfrak{y}\mathfrak{o} = \mathfrak{x}$ muß dies überdies eine der Klassen $M_1, M_2, \ldots, M_r$ sein. Da $\mathfrak{y}$ ebenfalls die Norm m hat, so liefert Satz 1 die Behauptung.

Aus Satz 24 folgt sofort der

Satz 31. *Jedes zum Führer k von* $\mathfrak{n}$ *teilerfremde Ideal in* $\mathfrak{o}$ *ist Erweiterungsideal.*

Definition 10. *Unter einer* ***Erweiterungsprimzahl*** *(in bezug auf die Diskriminante $D = k^2\Delta$) sei eine natürliche Primzahl p verstanden, die so beschaffen ist, daß jedes in p aufgehende Primideal des Körpers Δ Erweiterungsideal in bezug auf die Ordnung vom Führer k ist.*

Jede nicht in k aufgehende Primzahl ist Erweiterungsprimzahl; das ergibt sich aus Satz 31 unmittelbar. Aber es gibt auch Diskriminanten D, deren Erweiterungsprimzahlen teilweise in dem zugehörigen k aufgehen. Das zeigt folgender Satz:

Satz 32. *Ist die Primzahl p ein Teiler von k, aber nicht von Δ, und zugleich quadratischer Nichtrest nach Δ, d. h. im Körper Δ unzerlegbar, so ist das Primideal* $\mathfrak{p} = p\mathfrak{o}$ *Erweiterungsideal in bezug auf die Ordnung* $\mathfrak{n}$ *vom Führer k.*

Beweis. Es ist

$$\mathfrak{p} = (p\mathfrak{n})\mathfrak{o},$$

und $p\mathfrak{n}$ ist Ideal in $\mathfrak{n}$.

Definition 11. *Das Produkt aller in k aufgehenden Erweiterungsprimzahlen (mehrfache mehrfach gezählt) heiße der* ***größte Erweiterungsteiler*** *von k.*

Der größte Erweiterungsteiler von k hängt außer von k noch von $\varDelta$ ab. Satz 32 zeigt, daß er im allgemeinen von 1 verschieden ist.

Definition 12. *Eine Zahl m heißt zu k erweiterungsprim (in bezug auf die Stammdiskriminante $\varDelta$), wenn (m, k) Potenzteiler*[17]) *des größten Erweiterungsteilers von k ist.*

Satz 33. *Ist die natürliche Zahl m erweiterungsprim zu k, gibt es in der Idealklasse A aus der Hauptordnung ein Ideal $\mathfrak{x}$ von der Norm m und sind die Formenklassen $K_1, K_2, \ldots, K_r$ denjenigen Modulklassen von der Ordnung $\mathfrak{n}$ zugeordnet, die durch Multiplikation mit der Klasse O des Moduls $\mathfrak{o}$ die Klasse AO liefern, so ist m durch mindestens eine der Klassen K_i darstellbar.*

Beweis. Man zerlege $\mathfrak{x}$ in Primideale:

$$\mathfrak{x} = \mathfrak{p}_1 \mathfrak{p}_2 \ldots \mathfrak{p}_s.$$

Ist etwa p_i die durch $\mathfrak{p}_i$ teilbare natürliche Primzahl, so folgt

$$p_i / N(\mathfrak{x}),$$

$$p_i / m.$$

Entweder ist also p_i zu k teilerfremd, oder es ist $p_i/(m, k)$; im ersten Fall folgt aus Satz 31, im zweiten aus Definition 12 und 11, daß p_i Erweiterungsprimzahl und daher $\mathfrak{p}_i$ Erweiterungsideal ist. Offenbar gilt letzteres dann auch für das Produkt $\mathfrak{x}$ der $\mathfrak{p}_i$. Satz 30 ergibt also die Behauptung.

Auf Grund dieses Prinzips der Idealerweiterung ordnen sich schließlich auch die klassischen Ergebnisse über die Darstellung der zum Führer teilerfremden natürlichen Zahlen ein. Nach Satz 1 waren für die Darstellbarkeit der natürlichen Zahl m maßgebend die Ideale $\mathfrak{c}$ in $\mathfrak{n}$, soweit sie die Ordnung $\mathfrak{n}$ und die Norm m haben. Ist nun m teilerfremd zu k, so sind diese Ideale gemäß Satz 24 und 25 ein-eindeutig den sämtlichen zu k teilerfremden Idealen $\mathfrak{a}$ in der Hauptordnung zugeordnet vermöge der Beziehung

$$\mathfrak{c}\mathfrak{o} = \mathfrak{a} \tag{30}$$

oder der Beziehung

$$\mathfrak{c} = \mathfrak{a} \wedge \mathfrak{n}.$$

Im Anschluß hieran übertragen sich die Ergebnisse auf die Ideale der Hauptordnung, indessen erst dann, wenn auch die Klasseneinteilung übertragen ist. Eine Zuordnung der Modulklassen von der Ordnung $\mathfrak{n}$ zu den Idealklassen in der Hauptordnung ist wegen der im allgemeinen verschiedenen Anzahl nicht möglich. Wohl aber leisten die von H. Weber[29]) aus-

[29]) A. a. O., § 99.

führlich betrachteten und dort als „Idealklassen nach der Ordnung $\mathfrak{n}$“ bezeichneten Klassen das Gewünschte. Eine solche Klasse ist definiert als die Gesamtheit aller derjenigen Ideale $\mathfrak{b}$ in der Hauptordnung, die zu einem festen zu k teilerfremden Ideal $\mathfrak{a}$ in der Beziehung

$$\mathfrak{b} = \frac{\eta_1}{\eta_2}\,\mathfrak{a}$$

stehen, wo η_1 und η_2 ganze zu k teilerfremde Zahlen in der Ordnung $\mathfrak{n}$ sind und $\frac{\eta_1}{\eta_2}$ positive Norm hat[30]). Zu einer Darstellung einer zu k teilerfremden natürlichen Zahl m konstruiere man nun im Sinne des $\mathfrak{m}$-Algorithmus das Ideal $\mathfrak{c}$, welches ehemals die der Klasse K der darstellenden Form zugeordnete Modulklasse M lieferte. Das Ideal

$$\mathfrak{a} = \mathfrak{c}\mathfrak{o}$$

in der Hauptordnung ist zu k teilerfremd und hat die Norm m. Diejenige Idealklasse M' „nach der Ordnung $\mathfrak{n}$“, die das Ideal $\mathfrak{a}$ enthält, kann man nun im Anschluß an H. Weber[31]) der gegebenen Darstellung von m zuordnen. Aus dem erwähnten durch (30) vermittelten ein-eindeutigen Entsprechen folgt aber unschwer, daß die Zuordnung $M \rightarrow M'$ eine ein-eindeutige Zuordnung der Modulklassen von der Ordnung $\mathfrak{n}$ zu den Idealklassen (in $\mathfrak{o}$) nach der Ordnung $\mathfrak{n}$ bedeutet. Die in Betracht kommenden Ideale in zwei einander zugeordneten Klassen stehen paarweise in der Beziehung (30). Daraus ersieht man sofort: Eine zu k teilerfremde natürliche Zahl m ist dann und nur dann durch die Formenklasse K darstellbar, wenn es in der zugehörigen Klasse M' ein Ideal in $\mathfrak{o}$ von der Norm m gibt. Die klassische Deutung der Darstellbarkeit erscheint damit als ein nachträglicher Übergang zur Hauptordnung und ist, wie man sieht, streng an den teilerfremden Fall gebunden.

---

[30]) Im vorliegenden Spezialfall des quadratischen Körpers kann diese Klasseneinteilung auch charakterisiert werden als eine solche, die durch „Komplexion“ aus der bekannten, in der Klassenkörpertheorie üblichen Einteilung in „Strahlklassen“ entsteht. Zu diesen Begriffen vgl. H. Hasse, Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, Teil I: Klassenkörpertheorie (Jahresbericht der Deutschen Mathematiker-Vereinigung **35** (1926), S. 1—55, dort S. 5f. und 41). — Ein entsprechender Satz für Körper höheren Grades besteht nicht.

[31]) Dieser nimmt allerdings (a. a. O., § 99) die Klasse des Ideals $\mathfrak{a}$.

(Eingegangen am 26. 6. 1929.)

Der Dimensionsbegriff für Punktmengen in kartesischen Räumen.

Von

S. Bochner in München.

Inhaltsverzeichnis.

§ 1.
Problemstellung.

Die Dimensionstheorie der kartesischen Räume $\mathfrak{R}_n$ [1]) fing an mit dem folgenden Satz [2]).

Satz von Brouwer. *In einem $\mathfrak{R}_n$ ist es nicht möglich, eine Punktmenge, welche einen inneren Punkt enthält, durch eine topologische* [3]) *Abbildung in eine Menge ohne inneren Punkt überzuführen.*

---

[1]) Wegen der Begriffe: kartesischer Raum, Teilraum, Simplex, Komplex usw., verweisen wir auf den Anhang, in welchem der Vollständigkeit wegen verschiedene mehr oder minder geläufige Begriffe und Hilfsbetrachtungen zusammengestellt sind.

[2]) Es gibt drei direkte Beweise dieses Satzes: a) L. E. J. Brouwer, Math. Annalen 70 (1911), S. 161–165; b) H. Lebesgue, Fund. Math. 2 (1921), S. 257–285; c) E. Sperner, Hamburger Abhandl. 6 (1928), S. 265–272. Der zuletzt genannte Beweis ist außerordentlich einfach.

[3]) Die Abbildung einer Punktmenge $\mathfrak{A}$ auf eine Punktmenge $\mathfrak{A}'$ heißt topologisch, wenn sie umkehrbar eindeutig und umkehrbar stetig ist. Zwei Punktmengen $\mathfrak{A}$ und $\mathfrak{A}'$, zwischen denen eine solche Abbildung möglich ist, heißen homöomorph.

Dieser Satz führt zum Dimensionsbegriff, und zwar ergibt er die Berechtigung, dem $\mathfrak{R}_n$ und jeder Menge des $\mathfrak{R}_n$, welche einen inneren Punkt enthält, die Dimensionszahl n beizulegen. Aber für eine folgerichtige Dimensionierung allgemeiner Mengen in kartesischen Räumen, z. B. einer „Kurve" in der Ebene, bietet der Satz von Brouwer keine unmittelbare Handhabe.

Wir wollen nun im folgenden zeigen, daß man im engen Anschluß an die Fragestellung und den Gedankengang des Brouwerschen Satzes und unter schärferer Ausnutzung seiner Beweiselemente auch für die allgemeinsten Mengen in kartesischen Räumen einen brauchbaren Dimensionsbegriff aufstellen kann. Dieser Dimensionsbegriff ist demjenigen vollständig äquivalent, welcher der Dimensionstheorie allgemeiner topologischer Räume zugrunde liegt[4]); auf die Äquivalenzfrage werden wir im Schlußparagraphen näher eingehen. Aber ganz unabhängig davon werden wir „direkt" den folgenden Satz beweisen, in welchem der Brouwersche Satz als Korollar enthalten ist.

Allgemeiner Satz. I. *Einer jeden nicht-leeren Menge eines* $\mathfrak{R}_n$ *kommt eine der Zahlen* $0, 1, 2, 3, \ldots$ *als ihre Dimension zu.*

II. *Ist die Menge* $\mathfrak{A}$ *ein Teil der Menge* $\mathfrak{B}$, *so ist* $\dim \mathfrak{A} \leqq \dim \mathfrak{B}$.

III. *Die Dimension ist eine topologische Invariante, d. h. ist eine Menge* $\mathfrak{A}$ *eines* $\mathfrak{R}_n$ *homöomorph*[5]) *mit einer Menge* $\mathfrak{A}'$ *eines* $\mathfrak{R}_{n'}$, $n \lesseqgtr n'$, *so ist* $\dim \mathfrak{A} = \dim \mathfrak{A}'$. *Insbesondere wenn* $\mathfrak{A}$ *eine Menge eines* $\mathfrak{R}_n$ *ist, und dieser* $\mathfrak{R}_n$ *als Teilraum eines Umbettungsraumes* $\mathfrak{R}_{n'}$ *aufgefaßt wird, so hat* $\mathfrak{A}$ *dieselbe Dimension als Menge des* $\mathfrak{R}_n$ *wie als Menge des* $\mathfrak{R}_{n'}$.

IV. *Jede Menge des* $\mathfrak{R}_n$, *welche einen inneren Punkt enthält, also insbesondere der* $\mathfrak{R}_n$ *selbst, ist* n*-dimensional.*

V. *Jede* n*-dimensionale Menge des* $\mathfrak{R}_n$ *enthält einen inneren Punkt.*

Da wir den Dimensionsbegriff ab ovo und allgemeingültig entwickeln wollen, werden wir das, was wir im Anhang, der üblichen Terminologie folgend, einen n*-dimensionalen* Raum (Simplex, Komplex) genannt haben, bis auf weiteres einen n*-gradigen* Raum (Simplex, Komplex) nennen. Im Verlauf unserer Überlegungen wird sich herausstellen, daß „Grad" und „Dimension" dieselbe Zahl sind, wie man es ja von jedem Dimensionsbegriff verlangen muß.

§ 2.

Die Dimensionszahl für kompakte Punktmengen.

Wir betrachten bis auf weiteres nur solche Punktmengen, welche *kompakt*, d. h. abgeschlossen und beschränkt sind.

---

[4]) Vgl. K. Menger, Dimensionstheorie, 1928, insbesondere Kap. VIII.

Es sei in einem $\mathfrak{R}_n$ eine solche Menge $\mathfrak{A}$ gegeben. Wir betrachten im $\mathfrak{R}_n$ endlich viele Teilräume gleichen Grades m, in deren Vereinigungsmenge $\mathfrak{B}$ die Menge $\mathfrak{A}$ enthalten ist. Solche Vereinigungsmengen gibt es immer: der $\mathfrak{R}_n$ selbst ist ein $\mathfrak{B}$. Es liegt nahe, das Minimum der Zahlen m als Dimension von $\mathfrak{A}$ zu erklären. Um aber für „gekrümmte" und „zipflige" Mengen keine zu große Dimensionszahl zu erhalten, wollen wir die Menge $\mathfrak{A}$ zuerst etwas „glätten". Eine solche „Glättung" kommt durch eine „kleine" stetige Verrückung zustande, bei welcher auch ausdrücklich zugelassen wird, daß verschiedene Punkte in einen gleichen zusammenrücken dürfen. Wir verstehen (mit Brouwer, loc. cit.) unter einer *ε-Deformation*, $\varepsilon > 0$, einer Menge $\mathfrak{A}$ des $\mathfrak{R}_n$ eine stetige Abbildung von $\mathfrak{A}$ auf eine andere Menge des $\mathfrak{R}_n$, bei welcher jeder Punkt um weniger als die (Euklidische) Entfernung ε aus seiner ursprünglichen Lage verschoben wird. Ist $\varepsilon_1 < \varepsilon_2$, so ist jede ε_1-Deformation auch eine ε_2-Deformation.

Definition. *Die kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ heiße (in bezug auf diesen $\mathfrak{R}_n$) μ-dimensional, wenn μ das Minimum der Zahlen m von der folgenden Beschaffenheit ist: zu jedem $\varepsilon > 0$ gibt es eine ε-Deformation von $\mathfrak{A}$ auf einen Teil einer Vereinigungsmenge von endlich vielen m-gradigen Teilräumen von $\mathfrak{R}_n$.*

Man sieht leicht ein, daß man zur selben Dimensionszahl kommt, wenn man den Passus „*Vereinigungsmenge von endlich vielen m-gradigen Teilräumen*" durch den Passus „*Vereinigungsmenge von endlich vielen (nicht notwendig gleichgradigen) Teilräumen von der maximalen Gradzahl m*" ersetzt. Daher ist auf Grund der im Anhang durchgeführten Überlegung (vgl. Anhang § 3, 5.) die obige Definition mit der folgenden für die Anwendung viel handlicheren Formulierung äquivalent.

Definition. *Die kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ heiße (in bezug auf diesen $\mathfrak{R}_n$) μ-dimensional, wenn μ das Minimum der Zahlen m ist, für welche es zu jedem $\varepsilon > 0$ eine ε-Deformation von $\mathfrak{A}$ auf einen Teil*[5]) *eines m-gradigen Komplexes gibt.*

§ 3.

Ihre wichtigsten Eigenschaften.

Jetzt werden wir den allgemeinen Satz beweisen.

Ad I und II. Diese Behauptungen ergeben sich unmittelbar aus der Definition.

---

[5]) Unter einem Teil eines Komplexes verstehen wir einen echten Teil des Komplexes bzw. den ganzen Komplex.

Ad III. Wir werden ein Schlußverfahren von Brouwer[6]) mit einer geringen Modifikation wiederholen. Es ist leicht einzusehen, daß es genügt, wenn wir folgendes beweisen. Gegeben seien eine kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ und eine homöomorphe Menge $\mathfrak{A}'$ eines $\mathfrak{R}_{n'}$. Wenn es zu jedem $\varepsilon' > 0$ eine ε'-Deformation von $\mathfrak{A}'$ auf einen Teil $\mathfrak{A}''$ eines Komplexes $\mathfrak{K}''$ vom Grade $\leqq m$ gibt, so gibt es zu jedem $\varepsilon > 0$ eine ε-Deformation von $\mathfrak{A}$ auf einen Teil $\mathfrak{B}$ eines Komplexes $\mathfrak{K}_1$ vom Grade $\leqq m$.

Wir gehen von einem festen $\varepsilon' > 0$ und einem dazugehörigen $\mathfrak{K}''$ aus und nehmen an, daß jeder Simplex von $\mathfrak{K}''$ einen Durchmesser $\leqq \varepsilon'$ hat (vgl. Anhang § 3, Satz). Wir bezeichnen die Eckpunkte von $\mathfrak{K}''$ mit

$$(1) \qquad P_1'', P_2'', \ldots, P_\varrho''$$

und ordnen in folgender Weise jedem Punkte P_ν'' einen Punkt Q_ν aus $\mathfrak{R}_n$ zu. Zuerst nehmen wir aus irgendeinem der Simplexe, für welche P_ν'' Eckpunkt ist, irgendeinen Punkt von $\mathfrak{A}''$ und nennen ihn P_ν'''. Es ist[7])

$$(2) \qquad E(P_\nu'', P_\nu''') < \varepsilon'.$$

Die Punkte P_ν''' brauchen nicht untereinander verschieden zu sein. Jedem Punkte P_ν''' ordnen wir in $\mathfrak{A}'$ einen Punkt P_ν' zu, für welchen

$$(3) \qquad E(P_\nu''', P_\nu') < \varepsilon',$$

in solcher Weise, daß zusammenfallenden Punkten P_λ''' und P_μ''' zusammenfallende Punkte P_λ' und P_μ' entsprechen. Dem Punkt P_ν' aus $\mathfrak{A}'$ entspricht vermöge der Homöomorphie ein bestimmter Punkt aus $\mathfrak{A}$. Dies ist der Punkt Q_ν. Wir wollen jetzt die Zuordnung der Punkte Q_ν zu den Punkten P_ν'' zu einer stetigen Abbildung des Komplexes $\mathfrak{K}''$ auf einen Komplex $\mathfrak{K}_1$ vom Grade $\leqq m$ erweitern. Wir greifen irgendeinen Simplex $\mathfrak{S}''$ von $\mathfrak{K}''$ mit den Ecken

$$(4) \qquad P_{\lambda_1}'', P_{\lambda_2}'', \ldots, P_{\lambda_\sigma}''$$

heraus, stellen seine Punkte durch die Formel

$$(5) \qquad P'' = t_1 P_{\lambda_1}'' + t_2 P_{\lambda_2}'' + \ldots + t_\sigma P_{\lambda_\sigma}''$$

dar (vgl. Anhang § 1, 3.) und lassen dem Punkte P'' den Punkt

$$(6) \qquad Q = t_1 Q_{\lambda_1} + t_2 Q_{\lambda_2} + \ldots + t_\sigma Q_{\lambda_\sigma}$$

entsprechen. Dadurch wird $\mathfrak{S}''$ auf den eventuell entarteten Simplex mit den Ecken $Q_{\lambda_1}, Q_{\lambda_2}, \ldots, Q_{\lambda_\sigma}$ stetig abgebildet. Liegt der Punkt P in verschiedenen $\mathfrak{S}''$, so sind seine aus den verschiedenen $\mathfrak{S}''$ herrührenden Bild-

---

[6]) Loc. cit. § 2.

[7]) Mit $E(P, Q)$ werden wir die Euklidische Entfernung der Punkte P und Q bezeichnen.

punkte zusammenfallend; dies ist eine Folge davon, daß die Simplexe von $\mathfrak{K}''$ regulär sind (Anhang § 1, 4.) und nur in vollen Seiten aneinanderstoßen. Die Abbildungen der verschiedenen $\mathfrak{S}''$ schließen sich daher zu einer stetigen Abbildung von $\mathfrak{K}''$ auf eine Punktmenge zusammen, die man auf Grund des Satzes im § 3 des Anhangs als einen Komplex $\mathfrak{K}_1$ vom Grade $\leqq m$ auffassen kann. Hierbei geht $\mathfrak{A}''$ in einen Teil $\mathfrak{B}$ von $\mathfrak{K}_1$ über. Die auf dem Wege über $\mathfrak{A}'$ und $\mathfrak{A}''$ zustande kommende Abbildung von $\mathfrak{A}$ auf $\mathfrak{B}$ ist eine (stetige) Deformation von $\mathfrak{A}$, von der wir nachweisen wollen, daß sie durch passende Wahl von ε' kleiner gemacht werden kann als die vorgeschriebene Zahl $\varepsilon > 0$.

Es sei $\varepsilon > 0$ gegeben. Wegen der Kompaktheit von $\mathfrak{A}'$ können wir ε' so bestimmen, daß zwei Punkten aus $\mathfrak{A}'$, deren Abstand $< 4\varepsilon'$ ist, in $\mathfrak{A}$ Punkte entsprechen, deren Abstand $< \frac{1}{2}\varepsilon$ ist. Dem Punkte P aus $\mathfrak{A}$ mögen in $\mathfrak{A}'$ und $\mathfrak{A}''$ die Punkte P' und P'' entsprechen. Wir nehmen einen $\mathfrak{S}''$, mit den Ecken (4), in welchem P'' gelegen ist. Es ist

$$E(P'', P''_{\lambda_s}) < \varepsilon' \qquad (s = 1, 2, \ldots, \sigma). \tag{7}$$

Aus

$$E(P', P'') < \varepsilon' \tag{8}$$

und (2), (3) und (7) folgt

$$E(P', P'_{\lambda_s}) < 4\varepsilon' \qquad (s = 1, 2, \ldots, \sigma).$$

Daher ist

$$E(P, Q_{\lambda_s}) < \frac{1}{2}\varepsilon \qquad (s = 1, 2, \ldots, \sigma). \tag{9}$$

Insbesondere hat man

$$E(Q_{\lambda_r}, Q_{\lambda_s}) < \frac{1}{2}\varepsilon \qquad (r, s = 1, 2, \ldots, \sigma).$$

Aus (6) folgt, daß das Bild von $\mathfrak{S}''$ einen Durchmesser $< \frac{1}{2}\varepsilon$ hat, daher ist

$$E(Q, Q_{\lambda_s}) < \frac{1}{2}\varepsilon \qquad (s = 1, 2, \ldots, \sigma). \tag{10}$$

Aus (9) und (10) erhält man endlich

$$E(P, Q) < \varepsilon.$$

Ad IV. Eine Menge des $\mathfrak{R}_n$, die einen inneren Punkt enthält, enthält insbesondere einen n-gradigen Würfel $\mathfrak{W}$. Nach bisher Bewiesenem genügt es nachzuweisen, daß man $\mathfrak{W}$ nicht um weniger als ein beliebiges $\varepsilon > 0$ auf den Teil eines Komplexes vom Grade $\leqq n-1$ deformieren kann. Dies folgt aber sofort aus dem Satz von Brouwer[8]), daß bei jeder Deformation von $\mathfrak{W}$ um weniger als die halbe Kantenlänge die deformierte

[8]) Loc. cit. § 1.

Punktmenge eine offene Teilmenge enthalten muß. Aber ein Komplex im $\mathfrak{R}_n$ vom Grade $\leq n-1$ enthält keine offene Teilmenge.

Der Beweis des eben herangezogenen Hilfssatzes von Brouwer ist nicht ganz einfach. Wir wollen daher noch ein anderes Schlußverfahren angeben. Neuerdings hat E. Sperner[9]) einen sehr einfachen und eleganten Beweis für den folgenden Satz von Lebesgue angegeben.

a) Einen m-gradigen *Simplex* kann man nicht zu jedem $\varepsilon > 0$ auf endlich viele abgeschlossene Teilmengen $\mathfrak{T}$ vom Durchmesser $< \varepsilon$ aufteilen, von denen je $m+1$ einen leeren Durchschnitt haben.

b) Einen m-gradigen *Komplex* $\mathfrak{K}$ kann man zu jedem $\varepsilon > 0$ auf endlich viele abgeschlossene Teilmengen $\mathfrak{T}$ vom Durchmesser $< \varepsilon$ aufteilen, von denen je $m+2$ einen leeren Durchschnitt haben.

Zum Absatz b), dessen Formulierung etwas allgemeiner als bei Sperner[9]) ist, wollen wir die Konstruktion der $\mathfrak{T}$ bei gegebenem $\varepsilon > 0$ wiederholen. Wir nehmen an, daß die Durchmesser der Simplexe von $\mathfrak{K}$ weniger als $\frac{\varepsilon}{2}$ betragen. Wir greifen einen Simplex $\mathfrak{S}^{(\alpha)}$ von $\mathfrak{K}$ heraus. Sein Grad sei μ. Wir teilen $\mathfrak{S}^{(\alpha)}$, wie bei Sperner, auf $\mu+1$ abgeschlossene Teilmengen $\mathfrak{T}^{(\alpha)}$ auf, von denen jede einen der Eckpunkte, aber keinen Punkt der gegenüberliegenden Seite enthält. Diese Konstruktion führen wir für alle Simplexe aus, und vereinigen alle solche $\mathfrak{T}^{(\alpha)}$, welche einen Eckpunkt von $\mathfrak{K}$ gemeinsam haben zu *einer* Menge $\mathfrak{T}$. Jedes $\mathfrak{T}$ hat einen Durchmesser $< \varepsilon$ und enthält nur *einen* Eckpunkt von $\mathfrak{K}$. Ein Punkt P von $\mathfrak{K}$, der in einem μ-gradigen Simplex $\mathfrak{S}$ enthalten ist, kann höchstens in denjenigen $\mu+1$ Mengen $\mathfrak{T}$ vorkommen, welche Eckpunkte von $\mathfrak{S}$ enthalten.

Jetzt beweisen wir unsere Behauptung nach Alexandroff[10]) wie folgt. Die Menge $\mathfrak{A}$ des $\mathfrak{R}_n$ enthält einen n-gradigen Simplex $\mathfrak{B}$. Wäre dieser nicht n-dimensional, so könnte man ihn bei vorgegebenem ε auf einen Teil $\mathfrak{B}'$ eines höchstens $(n-1)$-gradigen Komplexes ε deformieren. Auf Grund von b) teilen wir $\mathfrak{B}'$ auf abgeschlossene Teilmengen $\mathfrak{T}'$ auf, deren Durchmesser $< \varepsilon$ ist, so daß der Durchschnitt von $n+1$ Mengen $\mathfrak{T}'$ leer ist. Zu jedem $\mathfrak{T}'$ ordnen wir in $\mathfrak{B}$ als Menge $\mathfrak{T}$ die Gesamtheit aller derjenigen Punkte zu, deren Deformationsbilder in $\mathfrak{T}'$ liegen. Der Durchmesser jedes $\mathfrak{T}$ ist kleiner als $\varepsilon_1 = 3\varepsilon$, überdies haben je $n+1$ Mengen $\mathfrak{T}$ einen leeren Durchschnitt. Da ε_1 beliebig klein sein kann, ergibt sich ein Widerspruch gegen a).

---

[9]) Loc. cit. § 4.

[10]) Math. Annalen **98** (1928), S. 617—636, insbesondere § 1 (und die inzwischen erschienene Arbeit in Annals of Math. (2) **30**, S. 101—187, insbesondere Kap. I. [Zusatz bei der Korrektur.])

Ad V. Wir brauchen nur zu zeigen: Jede kompakte Menge $\mathfrak{A}$ ohne innere Punkte hat eine Dimension $\leqq n-1$. Hierzu schließen wir $\mathfrak{A}$ in einen abgeschlossenen n-gradigen Würfel $\mathfrak{W}$ ein, und unterteilen $\mathfrak{W}$ in gleichkantige Würfel $\mathfrak{W}_1, \mathfrak{W}_2, \ldots, \mathfrak{W}_\varrho$ vom Durchmesser $< \varepsilon$. Im *Innern* eines jeden Würfels $\mathfrak{W}_\sigma$ greifen wir einen nicht zu $\mathfrak{A}$ gehörigen Punkt P_σ heraus. Einen solchen gibt es, da $\mathfrak{A}$ nach Voraussetzung keinen innern Punkt enthält. Von P_σ als Projektionszentrum projizieren wir jeden in $\mathfrak{W}_\sigma$ gelegenen Punkt von $\mathfrak{A}$ auf die Begrenzung von $\mathfrak{W}_\sigma$. Hierbei bleiben die Punkte auf der Begrenzung von $\mathfrak{W}_\sigma$ unverändert. Diese Projektionen ergeben zusammengenommen eine ε-Deformation von $\mathfrak{A}$ auf endlich viele $(n-1)$-dimensionale Teilräume. Die Dimension von $\mathfrak{A}$ ist also $\leqq n-1$.

Wir bemerken, daß das, was wir bisher den Grad eines Simplexes bzw. Komplexes genannt haben, nichts anderes als seine Dimension ist. Für einen Simplex folgt dies daraus, daß er, wenn er μ-gradig ist, als Punktmenge eines geeigneten $\mathfrak{R}_\mu$ aufgefaßt innere Punkte enthält. Für einen μ-gradigen Komplex folgt hieraus nach Behauptung II, daß er mindestens μ-dimensional ist; auf Grund der Definition ist er höchstens μ-dimensional, weil er eine 0-Deformation von sich selbst auf einen μ-gradigen Komplex ist. Also stimmen für einen Komplex der „Grad" und die „Dimension" überein. — Den oben zitierten Satz von Brouwer kann man folgendermaßen verallgemeinern. Eine m-dimensionale kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ kann nicht für jedes $\varepsilon > 0$ in eine niedrigerdimensionale Menge $\mathfrak{A}'$ ε-deformiert werden. Sonst brauchte man nur noch $\mathfrak{A}'$ seinerseits auf einen Teil eines Komplexes von einer Dimension $< m$ um weniger als ε zu deformieren, um (für jedes $2\varepsilon > 0$) eine 2ε-Deformation von $\mathfrak{A}$ auf einen Teil eines Komplexes von einer Dimension $< m$ zu erhalten.

§ 4.

Die Dimensionszahl für beliebige Punktmengen.

Nunmehr wollen wir die Dimension auch für eine nicht-kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ festsetzen. Wir gehen darauf aus, die Dimension von $\mathfrak{A}$ als die kleinste Dimensionszahl einer $\mathfrak{A}$ umhüllenden *kompakten* Menge zu definieren. Wollten wir es wörtlich bei der eben formulierten Dimensionsfestsetzung belassen, so würden sich sofort Unzuträglichkeiten einstellen. Erstens gibt es zu einer unbeschränkten Punktmenge keine beschränkte, also keine kompakte Hülle. Zweitens würden wir für jede überall dicht liegende Punktmenge des Intervalls $0 \leqq x \leqq 1$ die Dimensionszahl $\mu = 1$ erhalten, was aber für eine abzählbare lineare Punktmenge als eine „zu große" Dimensionszahl erscheinen würde. Und drittens wäre die Dimension keine topologische Invariante, weil man eine abzählbare lineare Punktmenge auf

den Teil einer kompakten linearen Punktmenge ohne inneren Punkt (das sogenannte Cantorsche Diskontinuum) topologisch abbilden kann, und eine solche Punktmenge nach § 2 die Dimension Null hat. Diese Unzuträglichkeiten scheiden aus, wenn wir folgendes festsetzen.

Definition. *Gegeben sei eine Menge* $\mathfrak{A}$ *in einem* $\mathfrak{R}_n$. *Wir betrachten irgendeine zu* $\mathfrak{A}$ *homöomorphe beschränkte Menge* $\mathfrak{A}'$ *in irgendeinem* $\mathfrak{R}_{n'}$ *und irgendeine kompakte Menge* $\mathfrak{B}'$ *aus diesem* $\mathfrak{R}_{n'}$, *welche* $\mathfrak{A}'$ *enthält. Unter der Dimension von* $\mathfrak{A}$ *verstehen wir das Minimum der Dimensionszahlen aller Mengen* $\mathfrak{B}'$.

Man sieht leicht, daß für eine kompakte Menge $\mathfrak{A}$ die „neue“ Dimension mit der „alten“ übereinstimmt. Diese Übereinstimmung beruht auf der (bereits nachgewiesenen) Gültigkeit der Behauptung III für kompakte Mengen.

Jetzt haben wir den allgemeinen Satz für beliebige Punktmengen nachzuweisen. Ad I ist nur daran zu erinnern, daß man den ganzen $\mathfrak{R}_n$ topologisch auf das Innere eines Würfels im $\mathfrak{R}_n$ abbilden kann, daß also jede Menge $\mathfrak{A}$ des $\mathfrak{R}_n$ jedenfalls auf eine beschränkte Menge (im $\mathfrak{R}_n$) abgebildet werden kann, daß also jeder Menge unsere Definition tatsächlich eine eindeutige Dimensionszahl zuordnet. Die Behauptungen II, III und IV ergeben sich unmittelbar. Die Behauptung V hingegen bedarf einer näheren Erörterung. Wir führen den Beweis wieder indirekt, indem wir nachweisen, daß jede Punktmenge $\mathfrak{A}$ des $\mathfrak{R}_n$, welche keinen inneren Punkt enthält, eine Dimension $\leqq n-1$ hat.

Zuerst nehmen wir an, daß $\mathfrak{A}$ nirgends dicht liegt, d. h. daß in der Umgebung eines jeden Punktes von $\mathfrak{A}$ nicht nur ein Punkt, sondern sogar eine n-dimensionale Kugel zur Komplementärmenge von $\mathfrak{A}$ gehört. Wir bilden den $\mathfrak{R}_n$ topologisch auf das Innere eines Würfels $\mathfrak{W}$ in einem neuen $\mathfrak{R}_n$ ab. Die Punktmenge $\mathfrak{A}$ geht hierbei in eine beschränkte nirgends dichte Punktmenge $\mathfrak{A}'$ des neuen $\mathfrak{R}_n$ über. Die kleinste abgeschlossene Hülle $\overline{\mathfrak{A}'}$ von $\mathfrak{A}'$ ist kompakt und wiederum nirgends dicht. Es ist

$$\dim \mathfrak{A} = \dim \mathfrak{A}' \leqq \dim \overline{\mathfrak{A}'} \leqq n-1.$$

Es sei nunmehr $\mathfrak{A}$ beliebig. Durch eventuelle Hinzunahme von Punkten des $\mathfrak{R}_n$ kann man erreichen, daß $\mathfrak{A}$ die Komplementärmenge einer im $\mathfrak{R}_n$ überall dicht liegenden *abzählbaren* Punktmenge $\mathfrak{C} \equiv (P_1, P_2, P_3, \ldots)$ ist. Zum Beweis unserer Behauptung genügt es offenbar, nachzuweisen, daß man eine solche Punktmenge $\mathfrak{A}$ topologisch auf eine nirgends dichte Punktmenge $\mathfrak{B}$ des $\mathfrak{R}_n$ abbilden kann. Solche Abbildungen kann man nach verschiedenen Prinzipien konstruieren; wir wollen *eine* solche Abbildung näher schildern.

Unsere Abbildung[11]) ist das Resultat von abzählbar vielen sukzessiven Abbildungen $\alpha_1, \alpha_2, \alpha_3, \ldots$, von denen jede den ganzen $\mathfrak{R}_n$ in sich selbst transformiert. Die durch α_1 hervorgerufenen Bilder von P, $\mathfrak{A}$, P_ν, $\mathfrak{C}$ bezeichnen wir mit P^1, $\mathfrak{A}^1$, P_ν^1, $\mathfrak{C}^1$, und allgemein die durch α_{m+1} hervorgerufenen Bilder von P^m, $\mathfrak{A}^m$, P_ν^m, $\mathfrak{C}^m$ mit P^{m+1}, $\mathfrak{A}^{m+1}$, P_ν^{m+1}, $\mathfrak{C}^{m+1}$. Bei der Abbildung α_1 entspricht jedem Punkt des $\mathfrak{R}_n$ genau ein Bildpunkt, nur dem Punkt P_1 entspricht eine ganze Punktmenge, die wir das „Loch" P_1^1 nennen werden. Bei der Abbildung α_2 entspricht jedem Punkt mit Ausnahme von P_2^1 je ein Punkt, also insbesondere dem Loch P_1^1 ein Loch P_1^2, nur dem Punkt P_2^1 entspricht ein ganzes Loch P_2^2. Allgemein sind die ersten m von den Gebilden $P_1^m, P_2^m, P_3^m, \ldots$ Löcher, die anderen Punkte. Die Abbildung α_m besteht im folgenden. Das dem Punkt P_m^{m-1} zugeordnete Loch P_m^m ist eine abgeschlossene n-dimensionale Kugel, deren Mittelpunkt in P_m^{m-1} gelegen ist, und deren Radius wir mit h_m bezeichnen. Und die Zuordnung von $\mathfrak{R}_n - P_m^{m-1}$ zu $\mathfrak{R}_n - P_m^m$ ist eine topologische Abbildung, die so zustande kommt, daß jeder Punkt P auf dem Strahl durch den Punkt P_m^{m-1} um die Entfernung h_m nach außen verschoben wird. Man rechnet leicht nach, daß hierbei die Entfernung für zwei Punkte, die nicht auf demselben Strahl liegen, vergrößert, also für irgend zwei Punkte jedenfalls nicht verkleinert wird. Von den Zahlen h_m setzen wir voraus, daß für $m = 1, 2, 3, \ldots$

$$r_m = h_{m+1} + h_{m+2} + \cdots < \frac{1}{2} h_m. \tag{11}$$

Wir greifen einen Punkt P^μ aus $\mathfrak{A}^\mu$ heraus, der von $P_{\mu+1}^\mu, P_{\mu+2}^\mu, \ldots$ verschieden ist, also bei den sukzessiven Abbildungen $\alpha_{\mu+1}, \alpha_{\mu+2}, \ldots$ nie in ein Loch verwandelt wird. Wegen

$$E(P^m, P^{m+k}) \leqq h_{m+1} + h_{m+2} + \cdots + h_{m+k} < r_m \qquad (m \geqq \mu) \tag{12}$$

ist die Folge P^m konvergent; den Grenzwert bezeichnen wir mit P^0. Es ist

$$E(P^m, P^0) \leqq r_m. \tag{13}$$

Aus

$$E(P^{m+1}, Q^{m+1}) \geqq E(P^m, Q^m)$$

folgt, daß zwei verschiedene Punkte P und Q aus $\mathfrak{A}$ zwei verschiedene Bildpunkte P^0 und Q^0 besitzen. Daher erhalten wir eine umkehrbar eindeutige Abbildung von $\mathfrak{A}$ auf eine Punktmenge $\mathfrak{A}^0$. Wir wollen zeigen, daß sie stetig ist. Es ist

$$\begin{aligned} E(P^0, Q^0) &\leqq E(P^0, P^m) + E(Q^0, Q^m) + E(P^m, Q^m) \\ &\leqq 2 r_m + E(P^m, Q^m). \end{aligned}$$

[11]) Diese Abbildung ist mir von Herrn C. Carathéodory freundlichst vorgeschlagen worden.

Zu gegebenem $\varepsilon > 0$ bestimmen wir ein m, für welches $2r_m < \frac{\varepsilon}{2}$, und bei festgehaltenem Index m und Punkte P bestimmen wir ein $\delta = \delta(\varepsilon)$, so daß aus

$$E(P, Q) < \delta$$

folgt $E(P^m, Q^m) < \frac{\varepsilon}{2}$ und daher

$$E(P^0, Q^0) < \varepsilon.$$

Hiermit ist die Stetigkeit der Abbildung bewiesen. Da bei der Abbildung die Punkte „auseinanderrücken", ist auch ihre Umkehrung stetig, also die Abbildung topologisch.

Jetzt beachten wir für festes m das Loch P_m^m. Wenden wir auf die Randpunkte P^m der Kugel P_m^m die Formel (12) an, so findet man, unter Berücksichtigung von (11), daß die Kugel vom Radius $\frac{1}{2}h_m$ um den Punkt P_m^{m-1} zu allen Löchern

$$P_m^m, P_m^{m+1}, P_m^{m+2}, \ldots$$

und daher auch zur Komplementärmenge $\mathfrak{C}^0$ von $\mathfrak{A}^0$ gehört. Es sei jetzt ein fester Punkt P^0 aus $\mathfrak{A}^0$ und die positive Zahl ε gegeben. Wir bestimmen ein festes μ, so daß

$$r_m < \frac{\varepsilon}{6} \quad \text{für} \quad m \geqq \mu.$$

Dann ist wegen (13)

$$E(P^0, P^\mu) < \frac{\varepsilon}{6}. \tag{14}$$

Nunmehr bestimmen wir ein m, so daß

$$E(P^\mu, P_{m+1}^\mu) < \frac{\varepsilon}{6} \quad \text{und} \quad m \geqq \mu. \tag{15}$$

Wegen (12) ist

$$E(P_{m+1}^\mu, P_{m+1}^m) < \frac{\varepsilon}{6}. \tag{16}$$

Die Relationen (14), (15) und (16) zusammengenommen ergeben

$$E(P_0, P_{m+1}^m) < \frac{\varepsilon}{2}.$$

Der Punkt P_{m+1}^m ist aber der Mittelpunkt einer zu $\mathfrak{C}^0$ gehörigen Kugel, also ist $\mathfrak{A}^0$ nirgends dicht im $\mathfrak{R}_n$.

§ 5.

Die Äquivalenzfrage.

Es erübrigt noch festzustellen, daß unsere Dimensionsdefinition mit der Urysohn-Mengerschen äquivalent ist. Das folgt für kompakte Mengen aus einem wichtigen Satz von Alexandroff[12]) und für allgemeinere Mengen

[12]) Loc. cit.

aus einem sehr schönen Satz von Menger[18]) in Verbindung mit einem Satz von Hurewicz[13]).

Der Satz von Alexandroff gab den Anstoß und die Grundlage zur vorliegenden Betrachtung. Er lautet folgendermaßen: Ist eine kompakte Menge $\mathfrak{A}$ eines $\mathfrak{R}_n$ μ-dimensional (im Urysohn-Mengerschen Sinne), so kann man $\mathfrak{A}$ für jedes $\varepsilon > 0$ in einen μ-dimensionalen Komplex, aber nicht für jedes $\varepsilon > 0$ in einen $(\mu - 1)$-dimensionalen Komplex ε-deformieren. Diese *Charakterisierung* der Urysohn-Mengerschen Dimension lautet fast ebenso wie unsere *Definition*. Der Unterschied besteht nur darin, daß bei Alexandroff ε-Deformationen auf *volle* Komplexe und bei uns auf *Teile* von Komplexen herangezogen werden. Aber dieser Unterschied ist nicht wesentlich. Denn im Grunde genommen beweist Alexandroff den Satz zuerst unter Zugrundelegung von ε-Deformationen auf Teile von Komplexen und verfeinert ihn hinterher in einer anschließenden Betrachtung zur endgültigen Formulierung. Genau genommen, d. h. dem Beweise nach, bezieht sich die Alexandroffsche Charakterisierung der Dimension nicht auf die Definition von Urysohn und Menger, sondern auf die nachfolgende Definition von Lebesgue: Die kompakte Menge $\mathfrak{A}$ ist μ-dimensional, wenn μ das Minimum der Zahlen m von der folgenden Beschaffenheit ist. Zu jedem $\varepsilon > 0$ kann man $\mathfrak{A}$ auf endlich viele abgeschlossene Punktmengen vom Durchmesser $< \varepsilon$ aufteilen, von denen je $m + 2$ einen leeren Durchschnitt haben. — Dieser Formulierung der Dimension ist ohne weiteres anzusehen, daß die Dimension topologisch invariant ist. Wir hätten daher den Beweis der Behauptung III für kompakte Mengen auch unter Benutzung des Alexandroffschen Satzes führen können.

Die Sätze a) von Hurewicz und b) von Menger lauten: a) Jeder separable metrische Raum endlicher Dimension ist Teil eines kompakten metrischen Raumes gleicher Dimension, b) jeder kompakte metrische Raum endlicher Dimension ist einer Menge eines kartesischen Raumes homöomorph.

Anhang über Simplexe und Komplexe.

§ 1.

Simplexe.

1. Unter dem n-dimensionalen (kartesischen) *Raum* $\mathfrak{R}_n$, $n = 1, 2, \ldots$, verstehen wir die mit der Euklidischen Metrik behaftete Gesamtheit aller Zahlen-n-tupel

$$\mathfrak{x} = (x_1, \ldots, x_n).$$

[18]) Menger, loc. cit., Kap. IX.

Zulässige Koordinatentransformationen sind Translationen und orthogonale Transformationen. Wir führen noch den $\mathfrak{R}_0$ ein: er besteht aus einem einzigen Punkt.

2. Unter einer *Linearform* $L(\mathfrak{x})$ des $\mathfrak{R}_n$ verstehen wir einen Ausdruck $l_0 + l_1 x_1 + \ldots + l_n x_n$ mit konstanten Koeffizienten l_ν. Ein $\mathfrak{R}_\mu$ ist *Teilraum* eines $\mathfrak{R}_n$, wenn er im $\mathfrak{R}_n$ durch ein System von Gleichungen

$$(1) \qquad L_\alpha(\mathfrak{x}) = 0 \qquad (\alpha = 1, 2, \ldots)$$

definiert ist. (Durch passende Koordinatenwahl kann man das Gleichungssystem (1), sofern es nicht identisch erfüllt ist, auf die Gestalt bringen:

$$x_{\mu+1} = x_{\mu+2} = \ldots = x_n = 0\,.)$$

Zu jeder Punktmenge $\mathfrak{A}$ im $\mathfrak{R}_n$ gibt es einen *„kleinsten"* umfassenden *Teilraum*. Wir bezeichnen ihn mit $\mathfrak{R}(\mathfrak{A})$.

3. In einem $\mathfrak{R}_n$ seien Punkte

$$(2) \qquad \mathfrak{x}_1, \mathfrak{x}_2, \ldots, \mathfrak{x}_{\mu+1} \qquad (\mu \geqq 0)$$

gegeben; mit t_ν bezeichnen wir reelle (skalare) Zahlen. Die Gesamtheit der Punkte

$$(3_1) \qquad \mathfrak{x} = t_1 \mathfrak{x}_1 + t_2 \mathfrak{x}_2 + \ldots + t_{\mu+1} \mathfrak{x}_{\mu+1},$$

$$(3_2) \qquad t_1 + \ldots + t_{\mu+1} = 1; \quad t_1, t_2, \ldots, t_{\mu+1} \geqq 0$$

nennen wir den *Simplex*

$$(4) \qquad \mathfrak{S} = \mathfrak{S}(\mathfrak{x}_1, \ldots, \mathfrak{x}_{\mu+1})$$

mit den Ecken (2). Jeder mit einem Teil der Ecken von $\mathfrak{S}$ gebildete Simplex, insbesondere jeder Eckpunkt von $\mathfrak{S}$, heiße eine *Seite* von $\mathfrak{S}$. Die nachfolgenden Behauptungen über Simplexe sind alle leichte Folgerungen aus der analytischen Formel $(3_{1,2})$.

4. Die Dimension des Teilraumes $\mathfrak{R}(\mathfrak{S}(\mathfrak{x}_1, \ldots, \mathfrak{x}_{\mu+1}))$, vgl. 2., ist $\leqq \mu$ Ist sie $= \mu$, so heiße der Simplex *regulär*, sonst *entartet*. Für einen regulären Simplex sind die Seiten ihrerseits reguläre Simplexe, und jeder Punkt des Simplexes wird durch die Werte der t_ν in *umkehrbar* eindeutiger Weise festgelegt.

5. Einen Simplex kann man rekursiv aus Strecken (d. h. zweieckigen Simplexen) aufbauen. Der $\mathfrak{S}_1 = \mathfrak{S}(\mathfrak{x}_1, \ldots, \mathfrak{x}_\mu)$ sei bereits konstruiert; hieraus entsteht der (4) als Vereinigungsmenge der Strecken, deren *ein* Eckpunkt in $\mathfrak{x}_{\mu+1}$ liegt und deren anderer Eckpunkt $\mathfrak{S}_1$ durchläuft.

6. Für einen regulären $\mathfrak{S}$ folgt hieraus — unter Berücksichtigung der Tatsache, daß $\mathfrak{x}_{\mu+1}$ nicht in $\mathfrak{R}(\mathfrak{S}_1)$ gelegen ist —, daß jeder Punkt von $\mathfrak{S}$, für welchen $t_\nu > 0$, $\nu = 1, 2, \ldots, \mu + 1$, ein *innerer* Punkt von $\mathfrak{S}$ in

bezug auf $\mathfrak{R}(\mathfrak{S})$ ist; hingegen ist jeder andere Punkt von $\mathfrak{S}$ Grenzpunkt (d. h. nicht-innerer Punkt) von $\mathfrak{S}$.

7. Wenn die Punkte $\mathfrak{x}_1', \ldots, \mathfrak{x}_{\mu+1}'$ eines $\mathfrak{R}_n$ bei einer senkrechten Projektion auf einen Teilraum des $\mathfrak{R}_n$ in die Punkte (2) übergehen, so projiziert sich der Simplex $\mathfrak{S}(\mathfrak{x}_1', \ldots, \mathfrak{x}_{\mu+1}')$ in den Simplex (4). Hieraus folgt leicht, daß man einen entarteten Simplex für $\mu \leq n$ (also für genügend großes n) in passender Weise als senkrechte Projektion eines regulären Simplexes auffassen kann. Da aber die (senkrechte) Projektion einer beschränkten abgeschlossenen Punktmenge identisch ist mit der Projektion ihrer Begrenzung, ist ein entarteter Simplex identisch mit der Vereinigungsmenge seiner Seiten. Hieraus findet man durch Schluß von μ auf $\mu+1$, wo $\mu+1$ die Eckenzahl bezeichnet:

Jeder entartete Simplex ist die Vereinigungsmenge von regulären Simplexen, nämlich die Vereinigungsmenge seiner regulären Seiten.

§ 2.

Konvexe Polyeder.

Wir betrachten nur abgeschlossene Punktmengen.

1. Eine Punktmenge $\mathfrak{A}$ des $\mathfrak{R}_n$ heißt konvex[14]), wenn mit irgend zwei Punkten $\mathfrak{x}_1$ und $\mathfrak{x}_2$ auch die Strecke mit den Eckpunkten $\mathfrak{x}_1$ und $\mathfrak{x}_2$ zu $\mathfrak{A}$ gehört. An der Darstellungsformel $(3_{1,2})$ erkennt man, daß jeder Simplex konvex ist. Aus § 1, 5. folgt, durch Schluß von μ auf $\mu+1$, daß mit irgendwelchen Punkten (2) auch der Simplex (4) zu $\mathfrak{A}$ gehört.

2. Durch Berücksichtigung von § 1, 6. erhält man hieraus unschwer, daß die konvexe Punktmenge $\mathfrak{A}$, als Punktmenge des $\mathfrak{R}(\mathfrak{A})$ betrachtet, innere Punkte enthält. Auf Grund dessen legen wir der Punktmenge $\mathfrak{A}$ eine Dimension bei, nämlich die Dimension des $\mathfrak{R}(\mathfrak{A})$.

3. Wenn man in einer *beschränkten* konvexen Punktmenge irgendeinen Punkt $\mathfrak{x}_0$, etwa einen inneren Punkt, festhält und alle Strecken betrachtet, die in $\mathfrak{x}_0$ anfangen und in einem Grenzpunkt von $\mathfrak{A}$ endigen, so ist ihre Vereinigungsmenge mit $\mathfrak{A}$ identisch.

Ist $\mathfrak{x}_0$ ein innerer Punkt und $\bar{\mathfrak{x}}$ ein Grenzpunkt von $\mathfrak{A}$, so sind die von $\bar{\mathfrak{x}}$ verschiedenen Punkte der Strecke von $\mathfrak{x}_0$ nach $\bar{\mathfrak{x}}$ innere Punkte von $\mathfrak{A}$.

4. Gegeben seien auf der Grenze einer konvexen Punktmenge zwei reguläre Simplexe $\mathfrak{S}_1$ und $\mathfrak{S}_2$, deren Durchschnitt $\mathfrak{S}_3$ eine volle Seite sowohl von $\mathfrak{S}_1$ als auch von $\mathfrak{S}_2$ ausmacht (und daher insbesondere ein Simplex ist), und ein Punkt $\mathfrak{x}_0$ im Innern dieser Punktmenge. Mit $\mathfrak{x}_0$ als weiterem Eckpunkt und mit der „Basis" $\mathfrak{S}_1$ bzw. $\mathfrak{S}_2$ bilde man gemäß § 1, 5. einen neuen

[14]) Zu den elementaren Eigenschaften konvexer Punktmengen vgl. C. Carathéodory, Rend. Circolo Matem. Palermo **32** (1911), S. 1—25, insbesondere Abschnitt I.

Simplex $\overline{\mathfrak{S}}_1$ bzw. $\overline{\mathfrak{S}}_2$. Der Durchschnitt von $\overline{\mathfrak{S}}_1$ und $\overline{\mathfrak{S}}_2$ — wir bezeichnen ihn mit $\overline{\mathfrak{S}}_3$ — füllt seinerseits je eine volle Seite von $\overline{\mathfrak{S}}_1$ und von $\overline{\mathfrak{S}}_2$ aus; er entsteht nämlich durch Hinzunahme des Eckpunktes $\mathfrak{x}_0$ zur „Basis" $\mathfrak{S}_3$.

5. In einem $\mathfrak{R}_n$ bilden wir mit *endlich vielen* Linearformen $L(\mathfrak{x})$ und $\Lambda(\mathfrak{x})$ das Relationssystem

$$(5_1) \qquad L_l(\mathfrak{x}) = 0 \qquad (l = 1, 2, \ldots),$$

$$(5_2) \qquad \Lambda_\lambda(\mathfrak{x}) \geqq 0 \qquad (\lambda = 1, 2, \ldots).$$

Die hierdurch definierte Punktmenge ist konvex. Wir nennen sie ein (konvexes) *Polyeder* $\mathfrak{P}$. Wir bemerken, daß das Relationssystem $(5_{1,2})$ in ein ähnlich beschaffenes übergeht, 1. wenn man im $\mathfrak{R}_n$ eine (zulässige) Koordinatentransformation vornimmt, 2. wenn man den $\mathfrak{R}_n$ in einen höherdimensionalen $\mathfrak{R}_m$ einbettet (es treten dann $m - n$ Relationen der Gestalt (5_1) hinzu), und 3. wenn man zu einem $\mathfrak{P}$ enthaltenden Teilraum $\mathfrak{R}_\mu$ von $\mathfrak{R}_n$ übergeht (es sind dann in einem passenden Koordinatensystem in den $L(\mathfrak{x})$ und $\Lambda(\mathfrak{x})$ die Koeffizienten von $n - \mu$ Koordinaten durch Nullen zu ersetzen).

6. Wenn man $\mathfrak{P}$ im $\mathfrak{R}(\mathfrak{P})$ betrachtet, so kommen die Gleichungen (5_1) in Wegfall[15]), weil solche Gleichungen unser Polyeder auf einen Teilraum von $\mathfrak{R}(\mathfrak{P})$ beschränken würden. Ersetzt man im allgemeinen System $(5_{1,2})$ für irgendein λ die Ungleichung $\Lambda_\lambda(\mathfrak{x}) \geqq 0$ durch $\Lambda_\lambda(\mathfrak{x}) = 0$, so entsteht eine Teilmenge von $\mathfrak{P}$, welche (sofern sie nicht leer ist) aus lauter Grenzpunkten von $\mathfrak{P}$ besteht und selber ein Polyeder, aber von niedrigerer Dimension als $\mathfrak{P}$ ist. Die Vereinigungsmenge dieser Polyeder umfaßt die gesamte Begrenzung von $\mathfrak{P}$. Denn jeder Punkt von $\mathfrak{P}$, für welchen

$$\Lambda_\lambda(\mathfrak{x}) > 0 \qquad (\lambda = 1, 2, \ldots),$$

ist ein innerer Punkt von $\mathfrak{P}$.

7. Jeder reguläre Simplex ist, wie man unschwer einsieht, ein Polyeder; desgleichen jedes Intervall (also insbesondere jeder Würfel), d. h. jede Punktmenge

$$a_\nu \leqq x_\nu \leqq b_\nu \qquad (\nu = 1, 2, \ldots).$$

8. Gegeben seien in einem $\mathfrak{R}_\mu$ endlich viele μ-dimensionale Polyeder $\mathfrak{P}_\alpha^*$ $(\alpha = 1, 2, \ldots)$ und endlich viele $(\mu - 1)$-dimensionale Teilräume $\mathfrak{R}_{\mu-1}^{(\beta)}$ $(\beta = 1, 2, \ldots)$. Das $\mathfrak{P}_\alpha^*$ werde durch das Relationssystem

$$L_\varrho^{(\alpha)}(\mathfrak{x}) \geqq 0 \qquad (\varrho = 1, 2, \ldots)$$

definiert, vgl. 6.; der $\mathfrak{R}_{\mu-1}^{(\beta)}$ durch die Relation

$$L_\beta(\mathfrak{x}) = 0.$$

---

[15]) Genauer genommen, wenn Formen $L(\mathfrak{x})$ auftreten, so sind sie identisch null.

Wir bezeichnen die Gesamtheit der Linearformen $L_\varrho^{(\alpha)}(\mathfrak{x})$ und $L_\beta(\mathfrak{x})$ für alle ϱ, α, β mit $\Phi_1(\mathfrak{x}), \Phi_2(\mathfrak{x}), \ldots, \Phi_k(\mathfrak{x})$, und wir betrachten das Relationssystem

$$\pm \Phi_\nu(\mathfrak{x}) \geqq 0 \qquad (\nu = 1, 2, \ldots, k)$$

für alle 2^k Kombinationen der positiven und negativen Vorzeichen. Auf diese Weise entstehen endlich viele Polyeder $\mathfrak{P}^{**}$, von denen wir zwei Eigenschaften feststellen wollen:

1) Falls ein solches Polyeder $\mathfrak{P}^{**}$ die Dimension μ besitzt (vgl. 2.), so kann es in seinem Innern weder einen Grenzpunkt eines andern $\mathfrak{P}^{**}$ noch einen Punkt eines $\mathfrak{R}_{\mu-1}^{(\beta)}$ enthalten, weil nämlich das Innere der $\mathfrak{P}^{**}$ die schärferen Relationen

$$\pm \Phi_\nu(\mathfrak{x}) > 0 \qquad (\nu = 1, 2, \ldots, k)$$

erfüllt (vgl. 6.).

2) Jedes $\mathfrak{P}_\alpha^*$ läßt sich als Vereinigungsmenge gewisser Polyeder $\mathfrak{P}^{**}$ auffassen, nämlich der Polyeder

$$L_\varrho^{(\alpha)}(\mathfrak{x}) \geqq 0 \qquad (\varrho = 1, 2, \ldots),$$
$$\pm \Psi_\tau^{(\alpha)}(\mathfrak{x}) \geqq 0 \qquad (\tau = 1, 2, \ldots),$$

wobei wir mit $\Psi_\tau^{(\alpha)}(\mathfrak{x})$ die von $L_\varrho^{(\alpha)}(\mathfrak{x})$ $(\varrho = 1, 2, \ldots)$ verschiedenen unter den Linearformen $\Phi_\nu(\mathfrak{x})$ bezeichnen.

§ 3.
Komplexe.

1. Gegeben seien in einem $\mathfrak{R}_n$ zwei Systeme von Polyedern, die Polyeder

$$(6) \qquad \mathfrak{P}_1, \mathfrak{P}_2, \ldots$$

und die Polyeder

$$(7) \qquad \mathfrak{P}_1^0, \mathfrak{P}_2^0, \ldots .$$

Wir sagen, daß die $\mathfrak{P}_\beta^0$ durch *Unterteilung* aus den $\mathfrak{P}_\alpha$ hervorgegangen sind, wenn jedes $\mathfrak{P}_\beta^0$ Teil eines $\mathfrak{P}_\alpha$ ist, und wenn man jedes $\mathfrak{P}_\alpha$ als Vereinigungsmenge von *endlich* vielen $\mathfrak{P}_\beta^0$ darstellen kann. Eine Unterteilung einer Unterteilung ist wiederum eine Unterteilung.

2. Satz. *Gegeben seien in einem $\mathfrak{R}_n$ endlich viele beschränkte (konvexe) Polyeder* (6).

Man kann sie in reguläre Simplexe

$$(8) \qquad \mathfrak{S}_1, \mathfrak{S}_2, \ldots$$

unterteilen, die folgenden zwei Bedingungen zugleich genügen.

A) *Der Durchschnitt irgend zweier Simplexe $\mathfrak{S}_\alpha$ und $\mathfrak{S}_\beta$ ist eine volle Seite sowohl von $\mathfrak{S}_\alpha$ als auch von $\mathfrak{S}_\beta$.*

B) *Der Durchmesser eines jeden Simplexes ist $< \eta$, wo η irgendeine fest gegebene positive Zahl bedeutet.*

Beweis. Die maximale Dimensionszahl der Polyeder (6) sei mit μ_0 bezeichnet. Diese Zahl ist, wie man leicht einsieht, eine Invariante des Polyedersystems (6) gegenüber Unterteilungen. Für $\mu_0 = 0$ sind alle $\mathfrak{P}_a$ Punkte, also unser Satz richtig. Wir nehmen an, daß er bereits für $\mu_0 \leqq \mu - 1$ bewiesen worden ist, und wollen ihn nunmehr für $\mu_0 = \mu$ beweisen. Den Übergang von (6) zu (8) werden wir in drei sukzessiven Unterteilungen vollziehen.

Erster Schritt. Wir betrachten im $\mathfrak{R}_n$ einen Würfel, der alle $\mathfrak{P}_a$ umschließt, unterteilen ihn in Würfel $\mathfrak{W}_\varrho$, deren jeder einen Durchmesser $< \eta$ hat, und bilden den Durchschnitt eines jeden $\mathfrak{P}_a$ mit einem jeden $\mathfrak{W}_\varrho$. Die Gesamtheit der resultierenden Polyeder, die nunmehr der Bedingung B) genügen, bezeichnen wir wiederum mit

(6) $$\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3, \ldots .$$

Zweiter Schritt. Wir konstruieren im $\mathfrak{R}_n$ endlich viele untereinander verschiedene Teilräume

(9) $$\mathfrak{R}_\mu^{(1)}, \mathfrak{R}_\mu^{(2)}, \ldots$$

von der Art, daß jedes μ-dimensionale unter den Polyedern (6) in einem $\mathfrak{R}_\mu^{(\nu)}$ enthalten ist. Wir bezeichnen: 1) mit $\mathfrak{R}_\mu$ einen festen unter den Teilräumen (9), 2) mit $\mathfrak{P}_a^*$ diejenigen μ-dimensionalen Polyeder aus (6), welche im $\mathfrak{R}_\mu$ enthalten sind, und 3) mit $\mathfrak{R}_{\mu-1}^{(\beta)}$ endlich viele Teilräume des $\mathfrak{R}_\mu$, in welchen die Durchschnitte der übrigen Polyeder (6) mit dem $\mathfrak{R}_\mu$ enthalten sind. Jetzt wenden wir die Konstruktion aus § 2, 8. an, welche jedes $\mathfrak{P}^*$ in neue Polyeder $\mathfrak{P}^{**}$ unterteilt. Wenn wir jetzt im Polyedersystem (6) jedes μ-dimensionale Polyeder in der angegebenen Weise unterteilen und in der neuen Polyedergesamtheit mehrmals vorkommende nur einmal zählen, so erhalten wir ein Polyedersystem — wir bezeichnen es wiederum mit

(6) $$\mathfrak{P}_1, \mathfrak{P}_2, \mathfrak{P}_3, \ldots \text{ —},$$

welches die folgende Eigenschaft besitzt:

E. Jedes μ-dimensionale Polyeder aus (6) enthält in seinem Innern keinen Punkt eines andern Polyeders aus (6).

Dritter Schritt. Diejenigen Polyeder aus (6), die eine Dimension $\leqq \mu - 1$ haben, bezeichnen wir mit $\mathfrak{P}''$, die anderen mit $\mathfrak{P}'$. Die Vereinigung der Begrenzungen aller $\mathfrak{P}'$ besteht gemäß § 2, 6. aus Polyedern $\overline{\mathfrak{P}}'$, deren Dimensionen $\leqq \mu - 1$ sind. Auf das aus den Polyedern $\mathfrak{P}''$ und $\overline{\mathfrak{P}}'$ bestehende System können wir nunmehr unsern Satz anwenden: Wir unterteilen diese Polyeder in Simplexe $\bar{\mathfrak{S}}$, welche die Eigenschaft A) besitzen. Das angestrebte Simplexsystem (8) bauen wir aus zwei Sorten von Simplexen auf, den Simplexen $\mathfrak{S}'$ und den Simplexen $\mathfrak{S}''$. Die Simplexe $\mathfrak{S}''$

sind alle diejenigen Simplexe $\tilde{\mathfrak{S}}$, welche Teil eines $\mathfrak{P}''$ sind. Die Simplexe $\mathfrak{S}'$ hingegen sind folgendermaßen definiert: Die Begrenzung eines festen $\mathfrak{P}'$ ist die Vereinigungsmenge von Simplexen $\tilde{\mathfrak{S}}$. Wir nehmen im Innern des $\mathfrak{P}'$ einen Punkt $\mathfrak{x}_0$ an und bilden gemäß § 1, 5. mit $\mathfrak{x}_0$ als Ecke und jedem der beteiligten $\tilde{\mathfrak{S}}$ als „Basis" einen Simplex $\mathfrak{S}'$. Die Vereinigungsmenge dieser $\mathfrak{S}'$ ergibt den $\mathfrak{P}'$, vgl. § 2, 3. Diese Konstruktion führen wir für alle $\mathfrak{P}'$ aus.

Zum Nachweis der Behauptung A) unterscheiden wir vier Fälle: 1) Die Simplexe $\mathfrak{S}_\alpha$ und $\mathfrak{S}_\beta$ sind beide vom Typus $\mathfrak{S}''$; 2) es ist $\mathfrak{S}_\alpha$ ein $\mathfrak{S}'$ und $\mathfrak{S}_\beta$ ein $\mathfrak{S}''$; 3) $\mathfrak{S}_\alpha$ und $\mathfrak{S}_\beta$ sind beide vom Typus $\mathfrak{S}'$ und gehören zum gleichen Polyeder $\mathfrak{P}'$ bzw. 4) zu verschiedenen Polyedern $\mathfrak{P}'$. Zum Fall 1) erinnern wir an die Definition der $\mathfrak{S}''$. Zum Fall 2) bemerken wir, daß (vgl. E. und § 2, 3.) der Durchschnitt von $\mathfrak{S}_\alpha$ mit $\mathfrak{S}_\beta$ gleich ist dem Durchschnitt der „Basis" von $\mathfrak{S}_\alpha$ mit $\mathfrak{S}_\beta$, also gleich ist dem Durchschnitt zweier $\tilde{\mathfrak{S}}$, und zum Fall 4), daß er gleich ist dem Durchschnitt der „Basis" von $\mathfrak{S}_\alpha$ mit der „Basis" von $\mathfrak{S}_\beta$. Endlich zum Fall 3) erinnern wir an § 2, 4.

3. Als sehr spezielle Folgerung aus unserm Satz stellen wir fest, daß man jedes (beschränkte) Intervall (vgl. § 2, 7.) als Vereinigungsmenge von endlich vielen regulären Simplexen vom Durchmesser $< \varepsilon$ darstellen kann.

4. Eine Vereinigungsmenge von endlich vielen regulären Simplexen (8), welche der Bedingung A) genügen, nennen wir einen *Komplex* $\mathfrak{K}$. Unter den Ecken von $\mathfrak{K}$ verstehen wir die Gesamtheit der untereinander verschiedenen Eckpunkte der Simplexe (8). Auf Grund unseres Satzes kann man eine Punktmenge, welche beschränkt und auf endlich viele Teilräume von der maximalen Dimension m verteilt ist, in einen m-dimensionalen Komplex einschließen.

(Eingegangen am 25. 5. 1929.)

Ein Knotensatz mit Anwendung auf die Dimensionstheorie.

Von

F. Frankl und L. Pontrjagin in Moskau.

Im folgenden wird bewiesen, daß *jedes* (und zwar auch jedes *verknotete*) einfach geschlossene Polygon des R^3 Rand einer singularitätenfreien orientierbaren Mannigfaltigkeit ist. Z. B. ist die Kleeblattschlinge Rand einer Fläche vom Geschlecht 1.

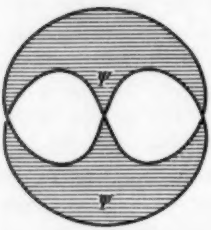
Mit Hilfe dieses Satzes wird dann gezeigt, daß für kompakte Teile des R^3 der Alexandroffsche Dimensionsbegriff[1]) mit dem Brouwer-Menger-Urysohnschen übereinstimmt[2]).

Wir beweisen zunächst den Satz über die Polygone. Sei also p ein einfach geschlossenes Polygon des R^3, so projizieren wir es von einem außerhalb gelegenen Punkt π, wodurch ein Kegel Φ entsteht. Den Punkt π wählen wir in allgemeiner Lage zu den Kanten und Eckpunkten des Polygons. Dann treten höchstens *Doppel*erzeugende auf. Wir triangulieren den Kegel so, daß jede Doppelerzeugende (d. h. die Strecke zwischen der Spitze und dem ersten Durchschnittspunkt mit dem Knoten) als Kante der Triangulierung auftritt. Die Dreiecke seien kohärent mit einem bestimmten Umlaufssinn des Knotens orientiert. Wir wollen nun die Fläche an den Doppelerzeugenden und an der Spitze so abändern, daß sie zu einer singularitätenfreien zweiseitigen Mannigfaltigkeit wird. Längs einer Doppelerzeugenden a stoßen genau vier Dreiecke zusammen. Wir numerieren sie entsprechend ihrer zyklischen Aufeinanderfolge im Raum mit $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ und zwar so, daß a in Δ_1 und Δ_2 entgegengesetzt orientiert erscheint (und natürlich ebenso in Δ_3 und Δ_4). Δ_1 und Δ_2 schließen den Winkelraum A, Δ_3 und Δ_4 den Winkelraum B ein. Wir verbinden nun den Mittelpunkt α

[1]) Alexandroff, „Zum allgemeinen Dimensionsproblem", Göttinger Nachrichten 6. Juli 1928, § 3, sowie „Untersuchungen über Gestalt und Lage abgeschlossener Mengen", Annals of Math. (2) **30**, S. 101—187, insbes. S. 183 u. f.

[2]) Der Satz über die Polygone und die beiden darauf folgenden Lemmata wurden von beiden Verfassern unabhängig voneinander gefunden. Im folgenden wird die Panklsche Form des Beweises wiedergegeben. Der dimensionstheoretische Satz stammt von Pontrjagin.

von a mit dem gegenüberliegenden Eckpunkt von Δ_i ($i = 1, 2, 3, 4$). Dadurch zerfällt Δ_i in zwei Dreiecke Δ_i' und Δ_i''. Wir nehmen jetzt in A in der Nähe von α einen Punkt α_A und ebenso in B einen Punkt α_B an und ersetzen die Dreiecke $\Delta_1', \Delta_2', \Delta_1'', \Delta_2''$ durch Dreiecke $\bar{\Delta}_1, \bar{\Delta}_2, \bar{\bar{\Delta}}_1, \bar{\bar{\Delta}}_2$, welche statt α den Eckpunkt α_A haben, und analog die Dreiecke $\Delta_3', \Delta_4', \Delta_3'', \Delta_4''$ durch Dreiecke $\bar{\Delta}_3, \bar{\Delta}_4, \bar{\bar{\Delta}}_3, \bar{\bar{\Delta}}_4$. Wenn α_A und α_B genügend nahe an α gewählt sind und wir diesen Prozeß an allen Doppelerzeugenden ausführen, so entsteht aus Φ eine orientierte Fläche Φ', in der jede Kante der Triangulierung höchstens zweimal vorkommt. Nur im Punkte π bleibt noch eine Singularität, welche folgendermaßen zu beheben ist: Wir legen um π eine kleine Kugel K. Diese schneidet Φ' in endlich vielen fremden einfach geschlossenen Linien, in welche man im Inneren von K fremde, einfach zusammenhängende Flächen einspannen kann. Ersetzen wir durch diese Flächen den im Inneren von K gelegenen Teil von Φ', so erhalten wir eine Fläche Ψ, welche alle Forderungen erfüllt.



Die nebenstehende Figur stellt eine nach der obigen Methode für die Kleeblattschlinge konstruierte Fläche dar.

Wir beweisen nun folgenden Hilfssatz: Sei M eine polyedrale orientierbare berandete zweidimensionale Mannigfaltigkeit im R^3 und r ihr Rand, dann gibt es beliebig viele ebensolche Mannigfaltigkeiten $\mathsf{M}_1, \mathsf{M}_2, \ldots, \mathsf{M}_n$ mit dem Rand r, welche aber sonst weder untereinander noch mit M Punkte gemein haben.

Wir zeigen zunächst, daß es in beliebiger Nähe zu M eine solche Mannigfaltigkeit M_1 mit dem Rand r gibt, welche sonst mit M keinen Punkt gemein hat. Wir zerlegen zu diesem Zwecke den R^3 simplizial so, daß M aus Dreiecken dieser Triangulation besteht und jedes Tetraeder höchstens *ein* (null-, ein- oder zweidimensionales) Randelement mit M gemein hat. Wir betrachten alle Tetraeder, welche mit M ein Randelement gemein haben, das nicht auf r liegt, und unter diesen alle, die auf einer vorgegebenen Seite von M liegen. Sei T ein solches Tetraeder und x das mit M gemeinsame Randelement; y sei das x gegenüberliegende Randelement von T. Wir ziehen alle Verbindungsstrecken von Punkten von x mit Punkten von y und betrachten die Menge ihrer Mittelpunkte; sie ist ein zweidimensionales Element. Alle so erhaltenen Elemente bilden zusammen eine Mannigfaltigkeit $\overline{\mathsf{M}}$, die M nicht schneidet[3]). Ihr Rand sei $\bar{r}$; er

[3]) Diese Methode der Approximation durch Mannigfaltigkeiten stammt von Alexandroff, Math. Annalen **98** (1928), S. 623, und Lefschetz, Ann. of Math. (2) **29** (1928), S. 241.

berandet zusammen mit r einen Komplex N, der aus einer Anzahl von zylindrischen Streifen besteht (jeder Randkurve von M entspricht einer). $\mathsf{M}_1 = \overline{\mathsf{M}} + \mathsf{N}$ hat dann die geforderte Eigenschaft. — Durch Wiederholung dieser Konstruktion gelangen wir zu beliebig vielen Flächen $\mathsf{M}_2, \mathsf{M}_3, \ldots, \mathsf{M}_n$.

Daraus folgern wir einen weiteren Hilfssatz, nämlich:

Sei Z ein Zylinder (d. h. die geschlossene Fläche, die aus dem Mantel und den beiden Basisflächen besteht). α_1 und α_2 seien die Endpunkte seiner Achse. Diese Endpunkte seien auf Z durch mehrere Linien $l_1, l_2, \ldots, l_n$ verbunden, wobei $l_i = r_{i1} + e_i + r_{i2}$, e_i eine Erzeugende des Zylinders und r_{1i} und r_{2i} Radien der beiden Basiskreise sind. Die Numerierung soll der zyklischen Aufeinanderfolge der Radien in den Basisflächen entsprechen. Außerdem seien α_1 und α_2 im Inneren von Z durch ein reguläres Polygon p verbunden. Dann gibt es im Inneren von Z zweiseitige Mannigfaltigkeiten M_i mit dem Rand $p + l_i$, welche bis auf Punkte von p zueinander fremd sind.

Beweis. Wir konstruieren zunächst nach den bisherigen Sätzen Flächen $\mathsf{N}_1, \mathsf{N}_2, \ldots, \mathsf{N}_n$ mit dem Rand $p + l_1$, welche sich sonst nicht schneiden; sie seien um die Achse in demselben Drehungssinn angeordnet wie die l_i; wie man leicht erkennt, kann man sie so wählen, daß sie im Inneren von Z liegen. Wir betrachten nun einen im Inneren von Z liegenden (etwa im bezug auf den Mittelpunkt homothetischen) etwas kleineren Zylinder Z'. Wir setzen $l_i' = \mathsf{N}_i \cdot \mathsf{Z}'$ und bezeichnen den im Inneren von Z' liegenden Teil von N_i mit N_i'. Aus Z' ragen zwei Strecken p_1 und p_2 von p heraus. In die geschlossenen Kurven $p_1 + l_i + p_2 + l_i'$ kann man nun paarweise fremde, zwischen Z und Z' gelegene orientierbare Flächen O_i einspannen[4]). Die Flächen $\mathsf{M}_i = \mathsf{N}_i' + \mathsf{O}_i$ erfüllen dann unsere Forderungen.

Wir gehen nun zum Beweis des dimensionstheoretischen Satzes über. Zu diesem Zwecke stützen wir uns auf Alexandroffs „Rechtfertigungssatz"[5]), dessen für uns in Frage kommender Spezialfall folgendermaßen lautet: Eine in sich kompakte Teilmenge F des R^3 ist dann und nur dann im

[4]) Dies zeigt man etwa so: Man kann annehmen, daß p_1 und p_2 in der Achse liegen, daß die Linien l_i' ebenfalls aus Radien und Erzeugenden bestehen und daß das zwischen Z' und Z gelegene Stück von $\mathsf{M}_1 = \mathsf{N}_1$ in einer von der Achse begrenzten Halbebene liegt Sei E_i die von der Achse begrenzte Halbebene, welche l_i trägt, und E_i' die von der Achse begrenzte Halbebene, welche l_i' trägt. Wir konstruieren nun ein System von zu Z homothetischen Zylindern $\mathsf{Z}_1 = \mathsf{Z}, \mathsf{Z}_2, \ldots, \mathsf{Z}_n = \mathsf{Z}'$, so daß Z_2 im Innern von Z_1, Z_3 im Innern von Z_2 liegt usw. Sei nun O_i' der Teil von E_i, der zwischen Z_i und Z liegt, O_i'' der Teil von E_i', der zwischen Z' und Z_i liegt, und O_i''' der Teil von Z_i, der im Winkelraum liegt, den man beschreibt, wenn man E_i' im angenommenen Drehungssinn um die Achse in die Lage von E_i dreht (wobei wir von der leicht realisierbaren Annahme ausgehen, daß die Winkel $\mathsf{E}_i \mathsf{E}_i'$ sehr klein sind). Dann können wir $\mathsf{O}_i = \mathsf{O}_i' + \mathsf{O}_i'' + \mathsf{O}_i'''$ setzen.

[5]) Alexandroff, Gött. Nachr., § 5; Annals, S. 184.

Sinne von Alexandroff höchstens eindimensional, wenn sie kein Gebiet des R^3 zerlegt. Da die Alexandroffsche Dimension niemals größer ist als die Brouwer-Menger-Urysohnsche[6]), müssen wir nur folgendes zeigen: Wenn F kein Gebiet des R^3 zerlegt, so ist F im Sinne von Brouwer, Menger und Urysohn höchstens eindimensional[7]).

Zum Beweise zerlegen wir den R^3 simplizial so, daß kein Eckpunkt dieser Triangulation in F enthalten ist. Da F in R^3 nirgendsdicht sein muß, gibt es beliebig feine Zerlegungen dieser Art. Sei k eine Kante dieser Zerlegung und S die Vereinigung aller Tetraeder mit der Kante k. Wir konstruieren nun einen Zylinder Z, dessen Achse Teil von k ist, der im Inneren von S liegt und den Durchschnitt $k \cdot F$ im Inneren enthält; dies machen wir für jede Kante und zwar so, daß alle so entstandenen Zylinder zueinander fremd werden. Nach dem Rechtfertigungssatz ist es nun möglich, die Endpunkte der Achse von Z im Inneren von Z durch ein reguläres Polygon p zu verbinden, welches zu F fremd ist. Seien nun $T_1, T_2, \ldots, T_n$ die Tetraeder mit der Kante k, entsprechend ihrer zyklischen Anordnung im Raume numeriert, und Δ_i die Dreiecke mit der Seite k (wobei $\Delta_1 = T_1 \cdot T_2$ usw.). Setzen wir $l_i = Z \cdot \Delta_i$, so können wir auf Z und die Linien $p + l_i$ das soeben bewiesene Lemma anwenden. Wir bekommen Flächen M_i, welche das Innere von Z, $J(Z)$ (in analoger Weise wie die Δ_i) in n Teile teilen, wobei an Stelle von $J(Z) \cdot T_i$ etwa M_i tritt. Wir ändern nun T_i so ab, daß $J(Z) \cdot T_i$ durch M_i ersetzt wird, und führen diesen Prozeß für alle Kanten und Tetraeder durch. An Stelle des Tetraeders T tritt dann eine dreidimensionale berandete Mannigfaltigkeit U. Stellen wir nun F als Summe aller $F \cdot U$ dar, so haben höchstens zwei dieser Teile Punkte gemein (nämlich, wenn die entsprechenden U eine Fläche gemein haben). Da diese Teile, indem man von einer genügend feinen Zerlegung des Raumes ausgeht, beliebig klein gemacht werden können, ist F nach dem Pflastersatz der Dimensionstheorie höchstens eindimensional im Sinne von Brouwer, Menger und Urysohn.

Zusatz bei der Korrektur (7. 10. 1929). Daß in dem obigen Beweis die Kurven $p + l_i$ im allgemeinen nicht unverknotet angenommen werden können, zeigt folgendes Beispiel, welches Alexandroff und Gružewski (Warschau) unabhängig voneinander gefunden haben.

Man geht aus von dem Würfel $0 \leq x \leq 1$, $0 \leq y \leq 1$, $0 \leq z \leq 1$ und bezeichnet ihn mit M_0. Diesen teilt man durch zu den Koordinatenebenen parallele Ebenen in $(2n-1)^3$ Teilwürfel, wobei n genügend groß

[6]) Annals, S. 183; Gött. Nachr., § 4.

[7]) Die Frage nach der Gültigkeit dieses Satzes wurde bereits von Urysohn, Fund. Math. 7, aufgeworfen.

gewählt sein muß. Nun läßt man aus M_0 einen verknoteten Kanal weg, der das mittlere Teilquadrat der oberen Basisfläche verbindet und aus Würfeln der Unterteilung besteht, so daß eine abgeschlossene Menge M_1 übrigbleibt (welche im wesentlichen mit dem abgeschlossenen Komplement eines verknoteten Torus homöomorph ist). In jedem Teilwürfel von M_1 wiederholen wir die Konstruktion, mit der Abänderung, daß jetzt der Kanal von links nach rechts geführt wird. Es entsteht eine Menge M_2. In jedem ihrer Teilwürfel wiederholen wir abermals die Konstruktion, nur daß jetzt die Kanäle von vorne nach hinten geführt werden. Dadurch entsteht eine Menge M_3, deren Teilwürfel wir wieder von oben nach unten durchbohren usw. Die Menge $M = \mathop{D}\limits_{i=0}^{\infty} M_i$ hat die geforderte Eigenschaft. Um zu beweisen, daß sie im klassischen Sinne eindimensional ist, muß man den oben bewiesenen Knotensatz wesentlich verwenden.

(Eingegangen am 8. 4. 1929.)

Über korrespondierende Punkte der Steinerschen Fläche vierter Ordnung und die Hauptpunkte derselben.

Von

Ólafur Daníelsson in Reykjavík (Island).

Es sei F eine Steinersche Fläche vierter Ordnung, T ihr Knotenpunkt, und d_1, d_2, d_3 ihre Doppelgeraden. Ein Punkt einer Doppelgeraden soll als zwei „einander gegenüberliegende Punkte" bezeichnet werden, deren je einer als jedem der beiden durch die Doppelgeraden gehenden Mäntel der Fläche angehörend aufgefaßt wird. Die Fläche hat vier singuläre Berührungsebenen, von welchen sie in den Punkten je eines Kegelschnittes berührt wird[1]). Diese Ebenen sollen die „Hauptebenen" der Fläche heißen, und die Berührungskegelschnitte werden als „Hauptkegelschnitte" bezeichnet.

Auf jeder Doppelgeraden gibt es zwei uniplanare Punkte, die mit ihren gegenüberliegenden Punkten zusammenfallen. Die Berührungsebenen der Fläche in diesen Punkten bilden ein vollständiges Vierkant, dessen Diagonalstrahlen die Doppelgeraden sind. Die vier Kanten schneiden die Fläche in Punkten, die ich die „Hauptpunkte" der Fläche nennen will. Es sollen nun die folgenden Sätze bewiesen werden:

1. Wenn X' ein Punkt der Fläche ist und man die Ebenen ξ_1, ξ_2 und ξ_3 durch X' und je eine der Doppelgeraden d_1, d_2 und d_3 legt, wird jede dieser Ebenen die Fläche in einem Punkte der Doppelgeraden berühren. *Die Berührungsebenen in den gegenüberliegenden Punkten gehen dann auch durch einen Punkt X'' der Fläche.* X' und X'' werden „korrespondierende Punkte" genannt.

2. Der Tangentenkegel vierter Ordnung eines Punktes der Fläche hat mit ihr die Berührungskurve gemein, und außerdem eine Kurve vierter Ordnung zweiter Art, die „Restkurve" des Punktes. *Die Restkurven zweier korrespondierenden Punkte liegen auf einer Fläche zweiter Ordnung.* Solche Kurven werden als „einander ergänzende Kurven" bezeichnet.

3. *Die Hauptpunkte der Fläche fallen mit ihren korrespondierenden Punkten zusammen. Jede Fläche zweiter Ordnung, die durch eine Haupttangentenkurve der Fläche geht, geht auch durch die vier Hauptpunkte der Fläche.*

4. *Die Restkurven der Hauptpunkte sind ebene Kurven vierter Ordnung mit drei Spitzen.* Ihre Ebenen werden „Nebenebenen" genannt.

5. *Die Berührungsebene in einem Hauptpunkte enthält die Schnittlinie der Nebenebene und der „zugehörigen" Hauptebene.*

Diese Sätze lassen sich mittels der gewöhnlichen Abbildung[1]) der Fläche leicht beweisen, indem man das Dreieck ABC (Fig. 1), auf welches die Doppelgeraden abgebildet werden, als Koordinatendreieck, und das Bild eines Hauptkegelschnittes als Einheitslinie ($x_1 + x_2 + x_3 = 0$) nimmt. Die vier Geraden $x_1 \pm x_2 \pm x_3 = 0$ (die Geraden f, g, h und k, Fig. 1) stellen dann die vier Hauptkegelschnitte dar[1]). K ist der Einheitspunkt ($1:1:1$) und die Hauptpunkte der Fläche sind auf die Punkte $1 : \pm 1 : \pm 1$ (die Punkte F, G, H und K) abgebildet.

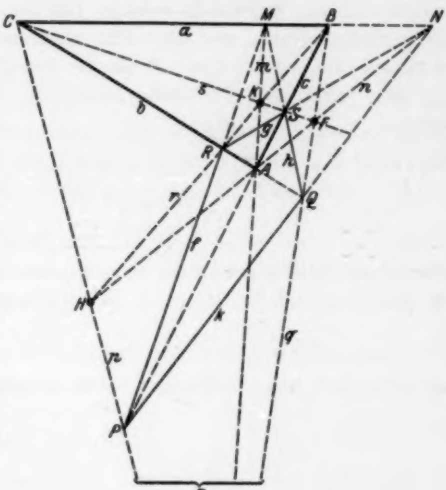


Fig. 1.

Die Punkte einer Koordinatenlinie (z. B. der Geraden a, mit der Gleichung $x_1 = 0$) sind einander paarweise zugeordnet und bilden somit eine Involution. Jedes Paar einer solchen Involution stellt einander gegenüberliegende Punkte der entsprechenden Doppelgeraden dar. Die Doppelpunkte dieser Involution sind die Punkte $0:1:1$ und $0:1:-1$, denn sie stellen die uniplanaren Punkte der Doppelgeraden d_1 dar, weil die Hauptkegelschnitte durch je drei der uniplanaren Punkte gehen. Ein solches Punktepaar hat somit die Koordinaten $0:1:t$ und $0:1:\frac{1}{t}$.

Jeder Kegelschnitt, der die Koordinatenlinien in solchen Paaren schneidet, muß eine ebene Kurve der Fläche darstellen. Ein solcher Kegelschnitt hat

[1]) Siehe z. B. Reye: Die Geometrie der Lage (1892). Dritte Abteilung, S. 146—152.

die Gleichung

$$\text{(I)} \qquad \alpha(x_1^2 + x_2^2 + x_3^2) + 2\alpha_1 x_2 x_3 + 2\alpha_2 x_3 x_1 + 2\alpha_3 x_1 x_2 = 0 \text{ [3]},$$

indem die Wurzeln der Gleichung

$$\alpha(x_2^2 + x_3^2) + 2\alpha_1 x_2 x_3 = 0$$

einander reziprok sind:

$$\frac{x_2}{x_3} = t, \frac{1}{t}.$$

Dasselbe Verfahren kann auf die Schnittkurve der Fläche mit einer Fläche zweiter Ordnung angewandt werden. Das Bild dieser Kurve soll eine Kurve vierter Ordnung sein, und diese Kurve soll durch neun Punkte vollständig bestimmt sein, und weiter soll sie die Koordinatenlinien in Punktepaaren der oben erwähnten Involution schneiden. Eine solche Kurve hat die Gleichung

$$\begin{aligned}\text{(II)} \quad \alpha(x_1^4 + x_2^4 + x_3^4) &+ \alpha_1(x_2^3 x_3 + x_3^3 x_2) + \alpha_2(x_3^3 x_1 + x_1^3 x_3) \\ &+ \alpha_3(x_1^3 x_2 + x_2^3 x_1) + \beta_1 x_2^2 x_3^2 + \beta_2 x_3^2 x_1^2 + \beta_3 x_1^2 x_2^2 \\ &+ \gamma_1 x_1^2 x_2 x_3 + \gamma_2 x_2^2 x_3 x_1 + \gamma_3 x_3^2 x_1 x_2 = 0,\end{aligned}$$

denn diese Gleichung enthält neun Konstanten, und die Kurve schneidet die Gerade $x_1 = 0$ in Punkten, deren Koordinaten durch die Gleichung

$$\alpha(x_2^4 + x_3^4) + \alpha_1(x_2^3 x_3 + x_3^3 x_2) + \beta_1 x_2^2 x_3^2 = 0$$

bestimmt sind, und die Wurzeln dieser Gleichung sind paarweise reziprok:

$$\frac{x_2}{x_3} = s,\ t,\ \frac{1}{s},\ \frac{1}{t}.$$

Der erste Satz leuchtet nun sofort ein. Es sei P' ein Punkt der Fläche und P_1 sein Bild in der Ebene (Fig. 2). Die Koordinaten von P_1 seien $p_1 : p_2 : p_3$. Weiter sei P'' der Punkt der Fläche, dessen Bild P_2 die Koordinaten $\frac{1}{p_1} : \frac{1}{p_2} : \frac{1}{p_3}$ hat. Legt man nun eine Ebene durch P' und die Doppelgerade d_1 (in a, mit der Gleichung $x_1 = 0$, abgebildet), so stellt die Gerade AP_1 die Schnittkurve dieser Ebene mit der Fläche dar. Die Gerade AP_2

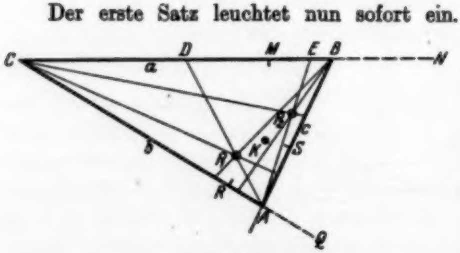


Fig. 2.

[3]) Salmon, A Treatise on the Analytic Geometry of three dimensions 2 (1915), S. 213.

dagegen ist das Bild der Schnittkurve der Fläche mit der Ebene $P''d_1$. Die Geraden AP_1 und AP_2 schneiden die Koordinatenlinie $x_1=0$ in den Punkten $0:p_2:p_3$ und $0:\frac{1}{p_2}:\frac{1}{p_3}$, und diese Punkte machen ein Paar der oben erwähnten Involution mit den Doppelpunkten $0:1:1$ und $0:1:-1$ aus. Die Ebenen $P'd_1$ und $P''d_1$ berühren daher die Fläche in einander gegenüberliegenden Punkten. P' und P'' sind dann korrespondierende Punkte.

Der allgemeine Kegelschnitt

$$\text{(III)}\quad \varphi_1 \equiv a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{23}x_2x_3 + 2a_{31}x_3x_1 + 2a_{12}x_1x_2 = 0$$

stellt eine Kurve vierter Ordnung zweiter Art dar, und der Kegelschnitt

$$\text{(IV)}\quad \varphi_2 \equiv \frac{1}{a_{11}}x_1^2 + \frac{1}{a_{22}}x_2^2 + \frac{1}{a_{33}}x_3^2 + 2\frac{a_{23}}{a_{22}a_{33}}x_2x_3 + 2\frac{a_{31}}{a_{33}a_{11}}x_3x_1 + 2\frac{a_{12}}{a_{11}a_{22}}x_1x_2 =$$

ihre ergänzende Kurve. Denn diese zwei Bildkurven schneiden die Koordinatenlinien in der oben erwähnten Involution. Siehe z. B. die Schnittpunkte der Kegelschnitte mit der Geraden $x_1=0$. Diese ergeben sich aus den Gleichungen

$$a_{22}x_2^2 + a_{33}x_3^2 + 2a_{23}x_2x_3 = 0$$

und

$$\frac{1}{a_{22}}x_2^2 + \frac{1}{a_{33}}x_3^2 + 2\frac{a_{23}}{a_{22}a_{33}}x_2x_3 = 0.$$

Die letztere dieser Gleichungen kann aus der ersteren erhalten werden, indem man x_2 mit $\frac{1}{x_2}$ und x_3 mit $\frac{1}{x_3}$ vertauscht, und die Behauptung ist bewiesen. Auch kann das Produkt $\varphi_1\varphi_2$ in der Form (II) geschrieben werden, was einen anderen Beweis der Behauptung liefert.

Das Bild eines ebenen Schnittes

$$\alpha(x_1^2+x_2^2+x_3^2) + 2\beta_1x_2x_3 + 2\beta_2x_3x_1 + 2\beta_3x_1x_2 = 0$$

zerfällt in die zwei Geraden

$$\alpha(a_1x_1+a_2x_2+a_3x_3)\left(\frac{x_1}{a_1}+\frac{x_2}{a_2}+\frac{x_3}{a_3}\right) = 0,$$

wenn die Ebene die Fläche berührt.

Die Gerade

$$\frac{x_1}{a_1}+\frac{x_2}{a_2}+\frac{x_3}{a_3} = 0$$

berührt im Punkte $a_1^2p_1 : a_2^2p_2 : a_3^2p_3$ den dem Koordinatendreieck eingeschriebenen Kegelschnitt

$$\sqrt{p_1x_1}+\sqrt{p_2x_2}+\sqrt{p_3x_3} = 0,$$

wenn $a_1p_1+a_2p_2+a_3p_3=0$, d. h. wenn die Gerade $a_1x_1+a_2x_2+a_3x_3=0$ durch den Punkt $p_1:p_2:p_3$ geht. Es sei P' der durch $p_1:p_2:p_3$ ab-

gebildete Punkt. Jede Berührungsebene durch P' berührt dann die in $\sqrt{p_1 x_1} + \sqrt{p_2 x_2} + \sqrt{p_3 x_3} = 0$ abgebildete Kurve. Diese Kurve ist demnach die Restkurve des Punktes P'. Sie ist vierter Ordnung, zweiter Art und berührt die Doppelgeraden der Fläche. Der im Punkte $\frac{1}{p_1} : \frac{1}{p_2} : \frac{1}{p_3}$ abgebildete zu P' korrespondierende Punkt P'' hat die in dem Kegelschnitt $\sqrt{\frac{x_1}{p_1}} + \sqrt{\frac{x_2}{p_2}} + \sqrt{\frac{x_3}{p_3}} = 0$ abgebildete Kurve zur Restkurve. Nun stellen aber die Kegelschnitte $\sqrt{p_1 x_1} + \sqrt{p_2 x_2} + \sqrt{p_3 x_3} = 0$ und $\sqrt{\frac{x_1}{p_1}} + \sqrt{\frac{x_2}{p_2}} + \sqrt{\frac{x_3}{p_3}} = 0$ zwei einander ergänzende Kurven dar, wie man sofort einsieht, indem man die Gleichungen auf die rationalen Formen:

$$p_1^2 x_1^2 + p_2^2 x_2^2 + p_3^2 x_3^2 - 2\, p_2 p_3 x_2 x_3 - 2\, p_3 p_1 x_3 x_1 - 2\, p_1 p_2 x_1 x_2 = 0$$

und

$$\frac{1}{p_1^2} x_1^2 + \frac{1}{p_2^2} x_2^2 + \frac{1}{p_3^2} x_3^2 - 2 \cdot \frac{1}{p_2 p_3} x_2 x_3 - 2 \cdot \frac{1}{p_3 p_1} x_3 x_1 - 2 \cdot \frac{1}{p_1 p_2} x_1 x_2 = 0$$

bringt, denn diese Gleichungen sind mit den Gleichungen (III) und (IV) identisch, wenn $a_{ii} = p_i^2$ und $a_{ik} = -p_i p_k$ ist $(i, k = 1, 2, 3)$. Die Gleichungen stellen somit einander ergänzende Kurven dar, und der zweite Satz ist bewiesen.

Die Punkte $p_1 : p_2 : p_3$ und $\frac{1}{p_1} : \frac{1}{p_2} : \frac{1}{p_3}$ fallen nur in den vier Punkten $1 : \pm 1 : \pm 1$ zusammen. Diese Punkte stellen die *Hauptpunkte* der Fläche dar.

Der Kegelschnitt

$$\frac{x_1^2}{k_1} + \frac{x_2^2}{k_2} + \frac{x_3^2}{k_3} = 0$$

ist, wenn $k_1 + k_2 + k_3 = 0$, dem Vierseit $x_1 \pm x_2 \pm x_3 = 0$ eingeschrieben, und stellt somit eine Haupttangentenkurve der Fläche dar. Die ergänzende Kurve der Haupttangentenkurve hat dann ihr Bild im Kegelschnitte

$$k_1 x_1^2 + k_2 x_2^2 + k_3 x_3^2 = 0.$$

Man hat nur in den Gleichungen (III) und (IV) $a_{ii} = \frac{1}{k_i}$ und $a_{ik} = 0$ $(i, k = 1, 2, 3)$ zu setzen. Ein solcher Kegelschnitt geht aber durch die Punkte $1 : \pm 1 : \pm 1$, weil $k_1 + k_2 + k_3 = 0$. Die Flächen zweiter Ordnung durch die Haupttangentenkurven gehen demnach auch durch die Hauptpunkte.

Wenn die Punkte $p_1 : p_2 : p_3$ und $\frac{1}{p_1} : \frac{1}{p_2} : \frac{1}{p_3}$ in den Punkten $1 : \pm 1 : \pm 1$ zusammenfallen, fallen ihre Restkurven in den ebenen Kurven

$$x_1^2 + x_2^2 + x_3^2 \pm 2\, x_2 x_3 \pm 2\, x_3 x_1 \pm 2\, x_1 x_2 = 0$$

— wo die Vorzeichen so zu wählen sind, daß das Produkt der drei letzten Glieder negativ ist — zusammen. Die die Restkurven enthaltenden Flächen zweiter Ordnung arten somit zu Doppelebenen aus. Diese „Nebenebenen" schneiden die Fläche in Kurven vierter Ordnung mit drei Spitzen. Nun enthält eine Hauptebene drei uniplanare Punkte der Doppelgeraden, und eine Nebenebene die drei übrigen. Sie sind „einander zugehörig". Eine Nebenebene wird von der zugehörigen Hauptebene in der Doppeltangente ihrer Schnittkurve geschnitten, weil jede Gerade einer Hauptebene eine Doppeltangente der Fläche ist. Die Schnittkurve der Nebenebene ist aber der Umriß der Fläche, wenn man den Hauptpunkt, dessen Restkurve die Schnittkurve ist, als Augenpunkt nimmt. Die Ebene durch den Augenpunkt und die Doppeltangente des Umrisses muß aber eine Berührungsebene der Fläche im Hauptpunkte sein, und damit ist der letzte Satz bewiesen.

Reykjavík, im Dezember 1928.

(Eingegangen am 1. 2. 1929.)

Erster Mathematiker-Kongreß
der Union der sozialistischen Räte-Republiken.

Die Mathematiker-Vereinigung der UdSRR veranstaltet den ersten Mathematiker-Kongreß der Sowjet-Union in Charkow (Ukraine) vom 24. bis 29. Juni 1930.

Alle Angehörigen der Sowjet-Union, die mathematische Wissenschaften treiben, haben das Recht am Kongreß teilzunehmen. Für die Mathematiker-Vereinigung der UdSRR wäre auch die Teilnahme ausländischer Mathematiker sehr wünschenswert.

Der Kongreß wird aus allgemeinen Versammlungen und Sektionssitzungen folgender 6 Sektionen bestehen:

1. Algebra und Zahlentheorie,
2. Funktionentheorie,
3. Differential- und Funktionalgleichungen,
4. Geometrie,
5. Mechanik und mathematische Physik,
6. Wahrscheinlichkeitsrechnung und mathematische Statistik.

Die Personen, die am Kongreß teilnehmen wollen, werden gebeten, sich wegen aller Auskünfte an das Organisationskomitee des Kongresses zu wenden (Charkow, Hauptpostamt, Postfach Nr. 333).

Das Organisationskomitee:

S. Bernstein, Vorsitzender,

W. Gontscharow, Sekretär für den Verkehr mit ausländischen Gelehrten.

721/7/57 — V/12/6